



BIG DATA AND AI FOR THE FINANCIAL SECTOR: CHALLENGES AND OPPORTUNITIES

V3 – Version 01.09.2022

**The Digital Finance ecosystems and transformative
trends towards future innovative research**

September 2022

TABLE OF CONTENTS

EXECUTIVE SUMMARY	5
1 Mission beyond Digital Finance	8
1.1 Introduction	8
1.2 Motivation	9
1.3 Objectives, vision and mission	10
2 Digital Finance Ecosystem	13
2.1 Market and trend analysis	14
2.2 The Challenges of Digital Transformation	21
2.3 Digital Transformation Areas in the Financial Sector	23
2.3.1 Digital Finance State of the Art Solutions	24
2.4 The impact of the COVID-19 Pandemic on the Financial Sector	27
3 Evolving Digital Finance towards 2025	29
3.1 Responsible Data-Driven AI	30
3.1.1 Data for AI	30
3.1.2 Security and Privacy	31
3.1.3 The evolving Regulatory Framework fostering EU competitiveness	34
3.1.4 Ethical AI in the Financial Sector	36
3.2 Deployment of applications and real-life scenario recommendations	43
3.3 The European Union market: the major AI-Fintech Hub of the world	46
4 Innovative Trends and Challenges in applying Big Data and AI in Digital Finance	49
4.1 AI and Robotics in Digital Finance	51
4.1.1 Ongoing challenges	52
4.1.2 EXplainable Artificial Intelligence (XAI)	54
4.2 Big Data Analytics in Digital Finance	58
4.2.1 Trends and Challenges	60
4.2.2 Emerging Big Data Solutions and examples of interesting applications	61
5 Recommendations towards future Research & Innovation Agendas	65
Annex I - Deep Dives on projects & applications	68
1.1 Digital Native Banks	68
1.2 Decentralized finance and comparative analysis of emerging cross chain services	71
1.3 Decentralized approaches to manage datasets used by AI and Big Data Systems	73

I.4 Network finance. A systemic and network approach designed to ease liquidation of trade receivables, optimise short-term credit allocation, and enhance and improve the efficiency of risk analysis models .. 74

Annex II Collection of use cases.....	77
<i>II.1 Banking Sector.....</i>	<i>77</i>
II.1.1 Intelligent Document Processing – IDP	77
II.1.2 Data Driven Decisions – 3D	79
<i>II.2 Financial Services.....</i>	<i>81</i>
II.2.1 Blockchain in the Digital Asset Supply Chain.....	81
II.2.2 DAOs in Decentralized Finance (DeFi).....	82
II.2.3 Network Finance	85
<i>II.3 Insurance Sector</i>	<i>87</i>
II.3.1 Personalized insurance products on IoT vehicles.....	87
II.3.2 Fraud detection and prevention in car insurances	88
<i>II.4 RegTech</i>	<i>90</i>
II.4.1 Automated Regulatory Risk Intelligence	90
II.4.2 Harmonised Compliance Protocol.....	92
<i>II.5 Others</i>	<i>94</i>
II.5.1 Audio Experience from Investors Reports.....	94

TABLE OF FIGURES

Figure 1 – “The New Physics of Financial Services” – World Economic Forum (2018)	9
Figure 2 – “Digital Finance Strategy for Europe”. Source: European Commission, 2020	12
Figure 3 - Number of Fintech start-ups worldwide from 2018 to February 2021, by region - BCG (FinTech Control Tower).....	15
Figure 4 - Global quarterly InsurTech investment activity - “Quarterly InsurTech Briefing”, Willis Tower Watson, 2021	16
Figure 5 – Global yearly investment activity by different financial technology sectors. Source: KPMG.	19
Figure 6 - Aggressive and conservative market estimates for autonomous systems and AI in the Fintech market – US market - Source: Autonomous Next	20
Figure 7 - Various segments of Autonomous Agents and Artificial Intelligence use in Fintech. Source: Autonomous Next	23
Figure 8 - Corporate operations in Fintech, 2014-2019, source: KPMG, Pulse of Fintech H2 2019	24
Figure 9 – “Insurtech private operations, 2014-2019” - Source: KPMG, Pulse of Fintech H2 2019	27
Figure 10 - Impact of the data-driven AI in the financial sector.....	29
Figure 11 - Data Center (INFINITECH Project)	44
Figure 12 – Testbeds (INFINITECH project).....	44
Figure 13 - Data Center, Testbed and Use Cases (INFINITECH Project)	45
Figure 14 - Financial Technology Space, Diagram by T. Butler	49
Figure 15 - Guide to understanding artificial intelligence , source: ABI Lab	51
Figure 16 - Application of AI, source: ABI Lab survey on ICT priorities.....	52
Figure 17 - Phase of an AI project, Andrew Yan-tak Ng	53
Figure 18 - Big Data Analytics applications	59
Figure 19 – Fintech trends and market value by type of service, source: TheBusinessResearchCompany 2021	66
Figure 20 - Digital Native Banks	69
Figure 21 - DeFi ecosystem survey across eight blockchain infrastructures.	72
Figure 22 - Data value chain hierarchy of needs	73

EXECUTIVE SUMMARY

Artificial Intelligence (AI) is not a recently discovered field. Since the beginning of the computer science discipline, in the late 1950s, AI has drawn a lot of attention in the international scientific community and since then it has represented a field of study which has triggered diverse and numerous research activities.

However, nowadays it seems that AI is getting a second wind. The extensive activities currently focused on this field can be explained by the technological maturity reached both in computational power (nowadays highly powerful hardware systems are available, small in size and with lower energy consumption), and in the real-time ability to analyse enormous quantities of data in any form.

Moreover, it is important to note how recently the phenomenon of **Big Data** has developed: the extensive availability and quantity of data, with volumes that have grown exponentially in the last few years, have contributed to a renewed interest in applications based on AI, also with a view to capitalising on the unexpressed potential of available data.

Currently, it is widely believed that the hype around AI can open a path for extensive transformations, not only in the economic and business spheres, but also socially and culturally, increasing the need to accompany the scientific and technological debate with ethical, regulatory and societal considerations.

Due to its importance in the development of the economic system, the financial sector is on the front lines in assessing and understanding how AI could, in the near future, fit into the reference ecosystem and favour **digital transformation processes**.

As with each innovating technology, the use of AI will bring with it not only new opportunities but also new risks, which give rise to new challenges in the regulatory arena: it is essential to consider how the products and services based on AI combine features of data dependency (data generation, processing, and analysis) with the connectivity, omnipresent in the new technological ecosystems.

In order for the huge attention surrounding AI to consolidate into a **real catalyst for the transformation of the financial sector**, it seems increasingly important to focus on the main ongoing challenges, which are, on the one hand, the internal operating factors (such as project management, the impact on risk mapping and cultural transformation) and on the other hand the factors inherent to the AI context (such as ethical aspects, regulatory environment and social impact).

In this overall landscape, the Big Data Value Association (BDVA) set up the **Task Force "AI and Big Data for the Financial Sector"**, led by GFT Italy and ABI Lab, with the purpose of investigating promising trends and emerging solutions across the EU through the adoption of an a holistic, ethically-sound and frictionless approach. Such approach allows to foster the innovation and competitiveness in the financial sector, while ensuring inclusiveness, protecting consumers and investors and addressing the emerging risks and challenges, so that the benefits of this fast-moving environment can be leveraged by all the actors involved.

This White Paper is exactly directed to illustrate the **major challenges and opportunities** related to AI and Big Data adoption in the financial sector and to proactively contribute to develop a common understanding and trajectory for the advancement of the financial technology (Fintech) landscape. This paper is aligned with BDVA Strategic Research,

Innovation and Deployment Agenda (SRIDA¹) and with the Digital Finance Package recently adopted by the European Commission (24 September 2020), especially regarding its “Digital Finance Strategy for Europe”.

The White Paper is structured as follows:

- the first section “**Mission beyond Digital Finance**” sets the scene and introduces the overall mission and objectives, motivation and vision;
- the second section “**Digital Finance Ecosystem**” describes the current market and trend analysis of the financial and insurance industry, focusing on several key topics, such as cybersecurity, augmented finance, payment, including digital banking, Property Technology (PropTech, i.e., technology for real estate), Insurance Technology (InsurTech), Regulatory Technology (RegTech) and a new generation of financial technology that create personal wealth management (WealthTech), which have as a common denominator the use of AI. This section also depicts the challenge of *digital transformation*, taking into account the current competitive landscape and the evolution of attitudes and demands and, more in general, the consumer perspective. Such challenges range from data fragmentation and interoperability barriers to regulatory constraints, to the data availability barrier, to the lack of validated business models as well as expected impact of new technologies on banking practices and business models. In addition, an overview of the key digital transformation areas in the domain is provided, describing the main State-of-the-Art solutions and services, and some remarks on the impact of the COVID-19 pandemic emergency in the financial sector;
- the third section “**Evolving Digital Finance Towards 2025**” contains insights on the evolution of the *Digital Finance*, which is expected to be data-driven and based on fast and reliable solutions, paving the ground for optimization and intelligent automation of its models, specifically analysing:
 - the responsible data-driven AI, focusing on the multi-facet importance of the data for training the AI systems;
 - the ethical AI in the financial sector and the evolving regulatory framework for scaling up of the Digital Finance in the EU. The most relevant recommendations and findings of the Expert Group on Regulatory Obstacles to Financial Innovation (ROFIEG)² are outlined, as well as reference is made to the Digital Finance Package, which are to move towards the creation of the European financial data space;
 - the Deployment of applications and real-life scenario recommendations, with special attention on the INFINITECH project ([Link to the site](#)), which is the flagship project in the financial sector;
 - the vision for enhance European positioning in the financial market at global level, starting from the exploitation of its role of largest Fintech laboratory in the world;
- the fourth section “**Innovative Trends and Challenges in applying Big Data and AI in Digital Finance**” provides an overview of the trends and challenges of applications of Digital Finance solutions, focusing on Big Data and AI and with the aim of moving ahead to bridge the gap between research and industry by showing the common

¹ https://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf

² https://ec.europa.eu/info/sites/default/files/business_economy_euro/banking_and_finance/documents/191113-report-expert-group-regulatory-obstacles-financial-innovation_en.pdf

interests and highlighting the synergies. In relation to AI and Robotics in Digital Finance the main ongoing challenges are described, followed by a focus on EXplainable Artificial Intelligence (XAI). For Big Data and Analytics, the analysis of the trends and analysis, in terms for instance of “siloeed” data systems, poor interoperability and issues related to the federated identity and access management, are accompanied by an overview of emerging Big Data solutions and examples of interesting applications, such as the Data Driven Credit Risk Scoring, Fast and Intelligent Anti-Money Laundering, the Robo-advisors, the Personalized Wealth Management, the Regulatory Sandboxes and the services for Regulatory Reporting and Compliance;

- the fifth section, relying on the hints captured during the analysis of the use cases, the existing and emerging applications, the challenges and trends, finally drafts a set of **“Recommendations towards future Research & Innovation”**.

The document also contains **two annexes**, prepared with consistent and high-quality inputs from the members of the BDVA Task Force: Annex I refers to the “Deep Dives on projects & applications” and Annex II consists of a collection of use cases, grouped in five categories (Banking Sector, Financial Services, Insurance Sector, RegTech and Others).

1 MISSION BEYOND DIGITAL FINANCE

1.1 INTRODUCTION

AI and Big Data provide the financial industry with valuable opportunities to increase efficiency and deliver high-quality services, thus representing the driving forces of Digital Finance. As a visionary perspective, in the short term, financial institutions can incorporate a whole variety of AI applications in their operating models to pursue new competitive strategies across their value chain, in mid-long term finance sector will transform the way they provide selected services to a more personalized experience with particular applications for an individual or small group of people. AI may have internal applications, which refer for example to the operational infrastructure of a bank, and key applications, which mainly target customer-oriented services. In this sense, both back-office and front-office operating models can enormously benefit from AI in terms of cost reduction, improved credit risk management, and an increase in collective solutions for shared problems.

As the portability of financial data will increase, we expect a shift in the competitive dynamics of the financial sector. Customer-oriented financial institutions will exert increasing market power and report growing profits. Their abilities to collect a large volume of customers' data and to derive more accurate decisions from AI-based processes represent valuable assets to their businesses. It is important to stress that AI and Big Data may have a positive impact not only on individual banks' business models but also on the overall financial system. Today, an increasing number of financial institutions are shifting towards an "open banking" environment, in which consumer banking, transactions, and other financial data are accessible to third-party financial services providers through the use of Application Programming Interfaces (APIs). Such collaborative initiatives built on shared data can improve the financial system's safety by increasing the timeliness, accuracy, and efficiency of non-competitive information.

It is expected that AI techniques and the development of computing methods will be increasingly embedded within the several architectural layers of the current financial technologies and thus they will create financial applications for the future of the local and global economies. Currently there are already some areas in finance and insurance (e.g., the creation of a new personalized investing portfolio based on financial transactions information, or the definition of insurance premiums based on behaviours of drivers respectively) where it is believed that AI applications immediately reflect the benefits. The current AI market is fragmented and does not have a common view towards where AI can benefit humans. However, in the financial sector, this common view is clearer. This is due to the fact that the finance instruments and the finance services are already well characterized with most implementations focusing on profit generation. Thereby the challenge to overcome is partially different and can be summarized as follow: in the financial sector it will be ensured that finance silo approaches and their transformation towards AI solutions will occur, and that AI and Big Data can be used also to strengthen security, safeguard assets, and reduce fraud.

Figure 1, generated by the World Economic Forum, depicts the six main areas for finance and also represents in tiny hexagons the different financial opportunities that exist per sector.

This paper aims at illustrating major challenges and opportunities related to AI and Big Data in the financial sector and at proactively contributing to develop a common understanding and trajectory for the advancement of the Fintech landscape towards 2025. The document lays down the foundations for cutting-edge European initiatives in the digital finance sector led by market players and the EC and in line with the EC strategy.

The Big Data Value Association (BDVA) set up a special task force, led by GFT Italy and ABI Lab, to investigate promising trends and emerging solutions across the EU. Through this task force, thanks to the adoption of a participatory process (actively involving and mobilizing the key stakeholders), the BDVA innovation ecosystem can be linked with the FinTech Market, with a holistic, ethically-sound and frictionless approach, capable of fostering the innovation and competitiveness in the financial sector, while ensuring inclusiveness and helping all the actors involved (EU economy, citizens and industry) to leverage the benefits of this fast-moving environment, as well as protecting consumers and investors, and properly addressing the emerging risks and challenges.

In accordance with the BDVA SRIDA to boost AI adoption in the EU, this White Paper looks after collecting different opinions and inputs from relevant stakeholders in their areas of expertise and financial sectors and it is the result of a joint effort and participation with multiple stakeholders' groups across the financial and insurance sector. In addition, this White Paper clusters a collection of different experiences and perspectives with the aim of shaping a consensus-built agreed-upon industry position and plan. As a result, this White Paper reflects the collective effort of a community that seeks to devise the next steps of the Fintech research agenda. In such collaborative ecosystem, the ultimate objective is to achieve a consensus on the role that AI and Big Data should play in the European financial sector.

1.3 OBJECTIVES, VISION AND MISSION

This White Paper aims at identifying relevant **Digital Finance trends** in the European context, and at contributing to set and consolidate a strong, human-centric and competitive European digital finance sector, thanks to the proposal of the **Digital Finance Vision 2025**, by establishing a common roadmap with clear steps for the short-/medium-period, meant to achieve the set objectives. These vision and roadmap are developed and implemented seeking continuous alignment with the **Digital Finance Package** recently adopted by the EU (24 September 2020) and especially its **“Digital Finance Strategy for Europe”**. The Strategy sets an overall picture of how the EU can strengthen the digital transformation of finance in the coming years, whilst minimizing the risks. The Digital Finance Package, which also includes the Retail Payments Strategies and legislative proposals on crypto-assets and digital operational resilience, is directed to encourage **responsible innovation in the EU financial sector** to benefit consumers and businesses, thereby protecting EU values. It also aims at mitigating potential risks related to money laundering, investor protection and cyber-crime, besides boosting the **EU competitiveness and innovation** in the financial sector and contributing to delivering the European Green Deal and the New Industrial Strategy for Europe. These objectives are the reference points guiding our White Paper development and implementation, as well as, in general, the BDVA Task Force on Digital Finance.

We acknowledge the words of **Valdis Dombrovskis** (2020, Press remarks by Executive Vice-President Valdis Dombrovskis on the “Capital Markets Union and Digital Finance”³),

³ https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_20_1739

Executive Vice-President For An Economy That Works For People, and are committed to work together, in the framework of the fertile BDVA environment, to contribute to their operationalization: *“The future of finance is digital. We saw during the lockdown how people were able to get access to financial services thanks to digital technologies such as online banking and fintech solutions. Technology has much more to offer consumers and businesses and we should embrace the digital transformation proactively, while mitigating any potential risks. That’s what today’s package aims to do. An innovative, strong and vibrant digital single market for finance will benefit Europeans and will be key to Europe’s economic recovery by offering better financial products for consumers and opening up new funding channels for companies.”*

To move in this direction, we are committed to promote a pan-European convergence and the exploitation of the potentials of a collaborative environment, where the key stakeholders work together under the BDVA umbrella to set the trajectories for future technological innovation in the Financial domain consistently with the Digital Finance Package, fostering the European Financial Market growth and upholding human well-being and empowerment through an ethically-aligned approach, thereby paving the way for influencing the Policy Makers towards a coherent regulatory development.

In this way, the Digital Finance vision 2025 will rely on the ecosystem driving the market evolution and will be capable of providing hints for policy development, based on promising concepts, and their expected positioning in the market, such as: Open Banking, e-money and evolving Anti-Money Laundering and Combating the Financing of Terrorism (AML/CFT) for crypto-assets, cyber-resilience, biometric authentication through fingerprint recognition, use of robo-advisors for investment advices, credit scoring facilitated by big data and machine learning, trade finance powered by Distributed Ledger Technology (DLT) and smart contracts, near-field communication through mobile wallet and the public cloud used for the outsourcing core banking/payment system, online crowdfunding platforms, and many others.

Accordingly, in this White Paper the Vision towards 2025 will be based on:

- current market trends, which has a high impact on business aspects;
- the ongoing regulatory landscape relevant to the shaping of suitable business scenarios;
- emerging technologies and related expected future market trends, paving the way for innovation beyond 2025, and giving rise to new business scenarios.

This task force provides a fertile environment where the different players can work together for the progress of a more competitive, transparent, and fair financial market, by developing and fully exploiting the potential of the financial ground-breaking innovations. The new and emerging financial technologies will be investigated and supported to ensure their uptake, capable of facilitating access to financial services and improving the efficiency of the financial system.

This White Paper and the Task Force efforts will be directed to move in synergy with key movements and initiatives, such as the mentioned Digital Finance Package, as well as the FinTech Action Plan (2018), the work of the European Parliament, European Supervisory Authorities (ESAs) and other experts, and the Digital Finance Outreach events and public consultations aimed at gathering feedback from a broad range of stakeholders.

In particular, this White Paper strongly relies on to the Digital Finance Strategy for Europe and the pillars on which it rotates around, i.e.: i) the European financial data space to promote data-driven innovation and finance; ii) the need to strengthen the Digital Single Market for financial services by removing its fragmentation; iii) address the risks and

challenges of digital transformation; iv) enhance the digital operational resilience of the financial system; v) the **central role of the EU regulatory framework for facilitating digital innovation** in the financial sector.

The key concepts of the Digital Finance Strategy that will be referred to alongside this document are:

- Enabling EU-wide interoperable digital identities in finance;
- Open Finance: promoting B2B data sharing in the financial sector;
- Clear and comprehensive European rules for crypto-assets;
- Mitigating the risk of digital transformation by strict and common rules on digital operational resilience;
- Ensuring “same activity, same risks, same rules”.



Figure 2 - “Digital Finance Strategy for Europe”. Source: European Commission, 2020 ⁴

⁴ https://ec.europa.eu/info/sites/default/files/business_economy_euro/banking_and_finance/200924-digital-finance-factsheet_en.pdf

2 DIGITAL FINANCE ECOSYSTEM

The Financial Stability Board (2017b) defines *Fintech* as the “technologically enabled financial innovations that could result in new business models, applications, processes, or products with an associated material effect on financial markets and institutions and on the provision of financial services”.

While innovation in finance is not a new concept, the focus on technological innovations and its pace have increased significantly. Fintech solutions that make use of Big Data analytics, AI, and blockchain technologies are currently introduced at an unprecedented rate. These new technologies are changing the nature of the financial industry, creating many opportunities that offer a more inclusive access to financial services. While the benefits of FinTech solutions are numerous, the introduction of new technological advancements in the financial sector also introduces some risks as well as concerns on consumer protection and financial stability. Relevant examples are underestimation of creditworthiness, market risk incompliance, fraud detection and cyber-attacks. Therefore, Fintech risk management represents a central point of interest for regulatory authorities and requires research and development of novel measurements (Giudici, 2018).

The EC stance on FinTech has been very clear and can be summarized as follows - EC acknowledges that digital finance has a lot to offer and believes that the people and businesses of the EU are ready for it. The EU must take full advantage of this in its recovery strategy to help repair the social and economic damage brought by the pandemic. Digital technologies will be key for relaunching and modernizing the European economy across multiple sectors. At the same time, users of financial services must be protected against risks stemming from increased reliance on digital finance. This stance is reflected in the EC Digital Finance Strategy for the EU (European Commission, 2020a) as well as in the EU budget powering the Recovery Plan for Europe (European Commission, 2020b). Having this in mind, the EU needs to close the gap between technical and regulatory expertise, in particular providing risk management procedures needed by both sides. This could lead to the development of a regulatory framework that encourages innovations in Big Data analytics, AI and blockchain technologies. At the same time, such regulatory framework needs to satisfy supervisory concerns in order to apply regulations in an effective and efficient way, while protecting consumers and investors.

Regulations and related supervisory requirements do place great focus on risk management practices, which in turn drive the need for deep, transparent and auditable data analyses across organizations. Technologies such as Big Data analytics, AI and blockchain ledgers may more efficiently address risk management requirements and the associated costs. In particular, these technologies can: (i) reduce credit scoring bias and improve fraud detection in peer-to-peer lending; (ii) measure and monitor systemic risk in peer-to-peer lending; (iii) measure and monitor market risk and volatility in financial markets; (iv) enhance client risk profile matching in robo-advisory; (v) identify illegal activities in crypto markets, including fraudulent initial coin offerings and money laundering; (vi) identify and prioritize IT operational risks and cyber risks.

2.1 MARKET AND TREND ANALYSIS

A recent biannual report by KPMG⁵ shows that total global Fintech investment remained high in 2019 with over \$135.7 billion invested globally across M&A (Merger & Acquisition), PE (Price to Earnings) and VC (Venture Capital) deals. Payments, including digital banking, remained the hottest area of Fintech investment globally, with a significant amount of focus on mature start-ups working to expand geographically or to grow their product breadth. Both PropTech and InsurTech investments were also quite strong in 2019, while RegTech saw a record number of deals despite a decline in total investment. At a technology level, cybersecurity-focused FinTech companies grew substantially on the radar of investors, while blockchain continued to attract a significant amount of attention (KPMG, 2020). According to this study, in 2019, investment in FinTech companies in Europe hit \$58.1 billion with 753 deals showing no signs of slowing down. On the other hand, European consumers have proven to be big fans of FinTech, further inspiring new approaches to the way financial services are delivered. A recent study by Ernst & Young (2019) has shown a significant increase in use of FinTech applications throughout Europe in the past years. Even though the need for different strategies around innovation and digital banking was apparent before the pandemic, COVID-19 has brought unprecedented disruption in the industry⁶. According to a Fintech Magazine's article⁷, the use of financial apps and mobile banking services has increased by 72% in Europe as a result of coronavirus. As depicted in Figure 3, this has also been reflected in the total number of FinTech start-ups worldwide which has almost doubled globally compared to 2019. In the remainder of the section, we will break down current FinTech market trends by different industry sectors according to recent studies.

5 <https://home.kpmg/xx/en/home/campaigns/2020/02/pulse-of-fintech-h2-19-global-trends.html>

6 https://www.ey.com/en_gl/financial-services-emeia/how-covid-19-has-spiced-up-digitization-for-the-banking-sector

7 <https://www.fintechmagazine.com/venture-capital/coronavirus-lockdown-sees-fintech-app-use-rise>

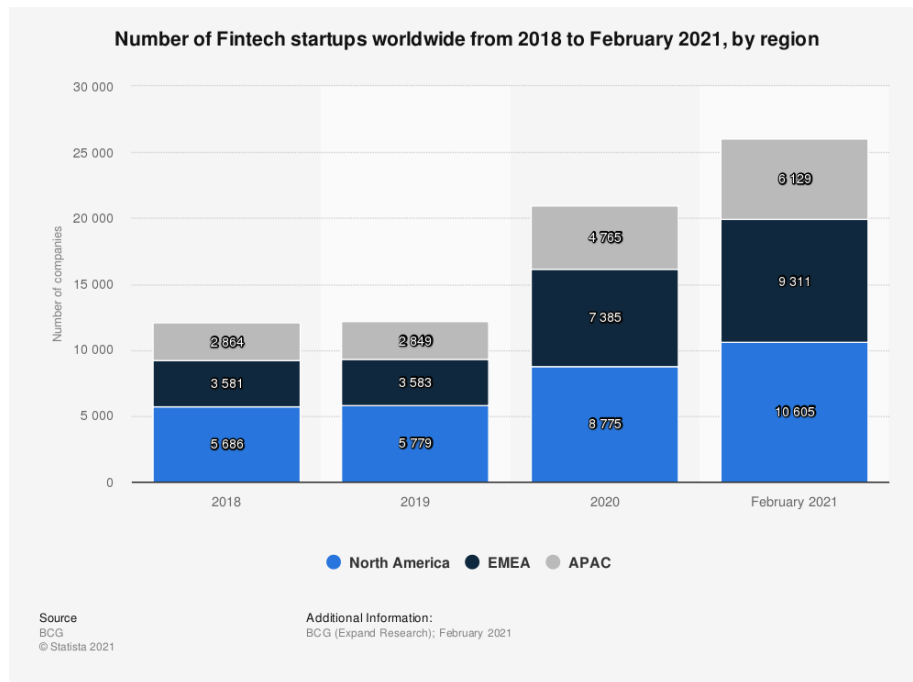


Figure 3 - Number of Fintech start-ups worldwide from 2018 to February 2021, by region - BCG (FinTech Control Tower).

Historically, the insurance services sector has been quite resistant to technology disruption and tended to favour those with deep pockets and long experience in the market. However, recent developments in Big Data and AI technologies have started to disrupt the sector thus coining the term InsurTech, which refers to a company using technology in order to disrupt the insurance industry. Large companies have taken notice and to stay competitive either choose to build up the knowledge themselves or by partnering with experienced research centres for data-driven business and AI⁸. In the EU market, Insurtech gained traction as the ecosystem continues to expand and grow. Hubs and innovation labs, like Insurtech Hub Munich and InsurLab Germany, continue to be instrumental in driving this growth (KPMG, 2020). Despite the COVID-19 outbreak, InsurTech funding finishes 2020 at a record high of \$7.1B across 377 deals, a 12% increase in funding and a 20% increase in deals compared to 2019⁹ (see Figure 4). For example, Lemonade¹⁰ became the first US InsurTech unicorn to go public and outside of the main hubs, thus global deal activity increased across several new geographies¹¹.

8 <https://www.know-center.tugraz.at/en/know-center-baut-mit-merkur-versicherung-die-digitale-infrastruktur-aus/>

9 <https://www.cbinsights.com/research/report/insurance-tech-q4-2020/>

10 <https://www.lemonade.com/de/en>

11 <https://www.cbinsights.com/research/report/insurance-tech-q2-2020/>

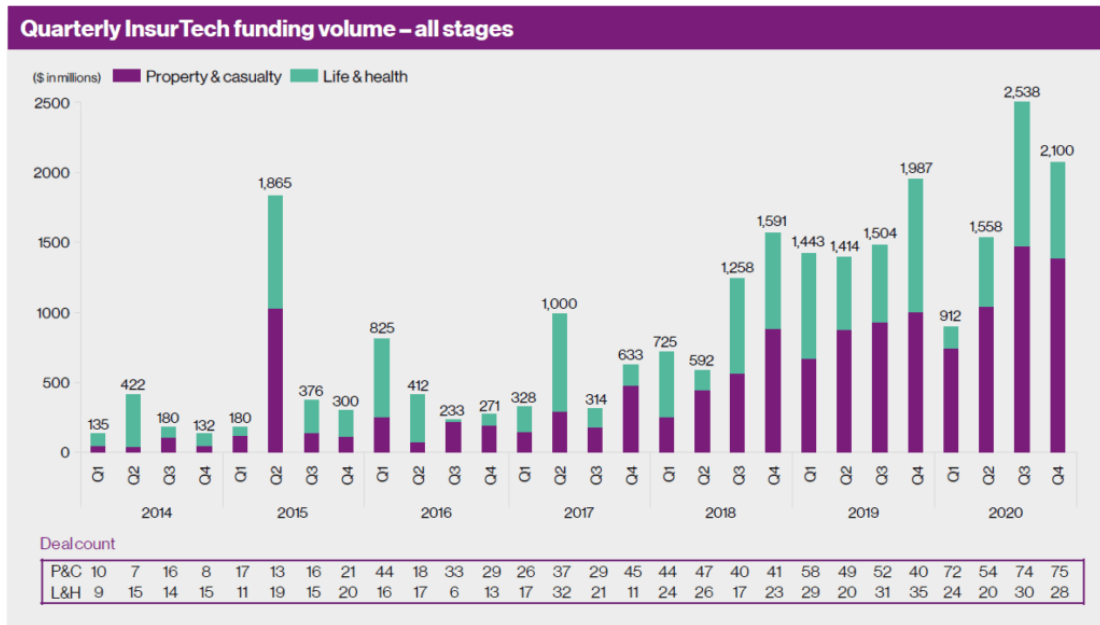


Figure 4 - Global quarterly InsurTech investment activity - “Quarterly InsurTech Briefing”, Willis Tower Watson, 2021¹²

We have also seen the rise of Cybersecurity investments (see Figure 4) thus becoming a hot topic, driven by the increasing importance of cybersecurity to both traditional financial institutions and FinTechs. Part of this relates to the move towards open banking, particularly in the UK and the EU. With data flows opening between different institutions, the ability to protect the data in transit or in the cloud is critical, not to mention protect it within institutions that are not regulated in the same way as financial institutions (KPMG, 2020). This has especially been a topic of interest in the EU with the introduction of the Payment Services Directive (PSD2). The key objectives of the PSD2 directive are to create a more integrated European payments market, and to make payments safer and more secure, thus protecting consumers. Another driver of cybersecurity investment over the past year has been the need for more effective and customer-centric access controls. Financial institutions are beginning to recognize the importance of providing an exceptional client experience. As a result, they are beginning to rethink their cumbersome identity management processes. This drives investment in FinTechs, offering innovative access management solutions such as biometrics and behavioural analytics (KPMG, 2020).

According to the KPMG research, despite a drop in deals volume from 2018 high, blockchain continued to be a hot topic in most regions of the world in 2019 with US and China dominating the funding¹³. The People’s Bank of China’s announcement of accelerated research and experimentation on digital currency and electronic payments have helped breathe new life into the space in 2019. The EU on the other hand provided funding for blockchain research and innovation through outright grants and prizes through the Horizon 2020 programme and supported investments. Significant budget for further grants is expected in the follow-up programme known as Horizon Europe¹⁴. Blockchain remained a hot topic for digital transformation and innovation until this day with Bitcoin reaching a record high price in February 2021 and an upcoming IPO (*Initial Public Offering*)

12 <https://www.willistowerswatson.com/-/media/WTW/Insights/2021/10/quarterly-insurtech-briefing-q3-2021.pdf?modified=20211026222035>

13 https://www.tagonline.org/wp-content/uploads/2020/05/CB-Insights_Blockchain-Report-2020.pdf

14 <https://ec.europa.eu/digital-single-market/en/blockchain-funding-and-investment>

of Coinbase (CB), the largest cryptocurrency exchange in the US with a potential \$100B valuation. However, according to the CB insights, governments are still uncertain around many questions that need to be answered with respect to digital currencies⁷. Furthermore, despite strong interest and greater awareness, blockchain has had minimal impact on the public sector, where few projects have moved beyond small pilots (Lindman et al., 2020)¹⁵.

Payments, including digital banking, remained the hottest area of Fintech investment globally, with a significant amount of focus on mature start-ups working to expand geographically or to grow their product breadth (KPMG, 2020). There has already been a significant decrease in cash usage over the past few years, but the COVID-19 pandemic caused an unprecedented surge in the demand for contactless payments. This has also led to outstanding performances for major companies offering cashless methods, such as Apple, Square and PayPal¹⁶ and there has been a number of significant M&A (*Merger & Acquisition*) deals in the payments space¹⁷. Furthermore, a number of large payment companies, namely PayPal, Square, Revolut, MasterCard and Visa have introduced or are planning to introduce payment services with cryptocurrencies¹⁸. According to KPMG, European digital banks (i.e., neobanks) showed their growing maturity, with many setting their sights on global expansion, including Revolut, N26, Starling, Tandem, Atom and others. European digital banks had over 25 million active users in 2020¹⁹. While several are targeting the US market, a number of digital banks in the UK are also looking to countries such as Australia, Hong Kong (SAR) and Singapore, while Germany-based N26 is working to expand to Brazil. Unfortunately, in 2020, a large scandal has shaken German FinTech scene: Wirecard, once a high-flying German electronic payments start-up, crashed into insolvency after disclosing that €1.9bn in corporate cash did not exist²⁰. Naturally, this has led the German government to make plans for tightening financial regulation laws. According to some analysts, Wirecard left a hole as big as \$30B in the European FinTech space and some European start-ups have already started to fill the void²¹.

PropTech is a term which refers to innovative technology solutions focused on real estate asset and property management and it has experienced a surge in investments in recent years. According to KPMG research, there are more real estate firms globally focusing on identifying innovation and PropTech opportunities and incorporating digital solutions within their primary business processes management. For example, companies like Zillow²² (US) and Rightmove²³ (UK) are already in the top listed companies in their respective markets. Although a relatively young field, PropTech start-ups have already raised over \$43B in funding worldwide since 2012. In 2018 alone there was an 82% increase when compared to the year before²⁴. When it comes to the EU, a research report from 2020²⁵ shows that the European PropTech is on the rise as sector funding grew 550% in five years. According to CNBC, the fastest growing trend in the housing industry is the demand for rental units

15 https://www.researchgate.net/publication/346036741_The_uncertain_promise_of_blockchain_for_government

16 <https://www.cnbc.com/2020/12/03/covid-19-pandemic-accelerating-the-shift-from-cash-to-digital-payments.html>

17 <https://www.fintechfutures.com/2020/12/top-10-fintech-and-bank-ma-deals-of-2020/>

18 <https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/for-paypal-visa-and-mastercard-cryptocurrencies-could-unlock-new-revenues-61239630>

19 <https://www.statista.com/statistics/941342/europe-largest-online-banks/>

20 <https://www.ft.com/content/2f5c9ce0-76f5-47cb-9835-fcc2524062b6>

21 <https://thenextweb.com/hardfork/2020/09/24/wirecard-fintech-startup-stock-dutch-adyen-mollie-amsterdam-fintech-week/>

22 <https://www.zillow.com/>

23 <https://www.rightmove.co.uk/>

24 <https://www2.deloitte.com/content/dam/Deloitte/pt/Documents/RealEstate/repredictions2020/deloitte-re-predictions-2020-5.pdf>

25 <https://tech.eu/brief/proptech-trends-proptech1-ventures-research-report/>

and for fast development of vertical rental properties, a trend also known as the “rent generation”. This shift in preference from owning to renting real estate is a major driver of the PropTech trend. Just as FinTech grew out of the 2007-08 economic downturn, experts say the COVID-19 pandemic gives PropTech companies an opportunity to innovate and offer new solutions in the real estate landscape and it is already attracting a lot of attention from entrepreneurs, investors, and industry incumbents all at once²⁶.

The term RegTech was first coined by the UK’s Financial Conduct Authority (FCA) in 2015 who called it: “A subset of FinTech that focuses on technologies that may facilitate the delivery of regulatory requirements more efficiently and effectively than existing capabilities.” In simple terms it refers to any technology that ensures companies comply with their regulatory requirements. An increasingly complex regulatory environment has led to an increase in compliance costs. Regulators have also issued heavy fines for non-compliance. As per CB Insights, firms have paid \$321B in fines since the 2008 financial crisis. In this context, firms are looking for solutions that help them reduce the risk of non-compliance. RegTech promises to disrupt the regulatory landscape by providing technologically advanced solutions to the ever-increasing demands of compliance within the financial industry²⁷. RegTech is gaining traction following the implementation of the European Parliament’s MiFID II legislation in January 2018. This regulation aimed to provide uniformity among investment services within the 31-member states of the European Economic Area (EEA). Several other regulations have also positively impacted the field such as Basel II, Solvency II, PSD2, the General Data Protection Regulation (GDPR), the Dodd-Frank Wall Street Reform and Consumer Protection Act²⁸. In 2020, the European Banking Authority (EBA) launched a RegTech industry survey to invite all relevant stakeholders, such as financial institutions and ICT third party providers, to share their views and experience on the use of RegTech solutions, on a best effort basis. The aim of the survey is to better understand the ongoing activity in this area, raise awareness on RegTech within the regulatory and supervisory community, and inform any relevant future policy discussion²⁹. There is already a number of EU companies and start-ups in the RegTech space and this number is only going to grow³⁰.

WealthTech is a term which refers to a new generation of financial technology companies that create digital solutions to transform the investment and asset management industry making the task of personal wealth management and growth now also in the hands of technology. In recent years, new companies have arrived on the scene offering advice based on AI and Big Data, micro-investment platforms, or trading solutions based on social networks³¹. According to a KPMG study, many WealthTech investors are focusing less on early-stage companies with interesting ideas and more on those companies in their portfolios looking for follow-on investments. Data management and analysis was a key focus for investment this year, with companies looking for more effective ways to assess and report on real-time data. Furthermore, financial advisors put increasing pressure on asset managers to create very specific products and investment solutions focused on the needs of niche groups of clients such as customers interested in making sustainable investments (KPMG, 2020). In 2018, a market intelligence platform Invyo has created a

26 <https://news.crunchbase.com/news/the-new-proptech-startups-focused-on-saving-time-and-money-will-thrive/>

27 <https://www2.deloitte.com/lu/en/pages/technology/articles/regtech-companies-compliance.html>

28 <https://www.eqs.com/compliance-blog/what-is-regtech/>

29 <https://www.eba.europa.eu/eba-consults-use-regtech-solutions-and-ways-support-uptake-regtech-across-eu>

30 <https://selleo.com/blog/top-25-regtech-companies-in-european-union-in-2020>

31 <https://www.bbva.com/en/what-is-wealthtech/>

mapping of WealthTech players in Europe³². It includes 100 of the most active players coming from 12 European countries that were categorized in six sub-segments: Investment tools, Data analytics, Portfolio management, Software, Digital brokerage and Robo advisors. WealthTech solutions allow also small investors to enter in the market benefitting of an automatized and professional service, thus lowering the barrier of entrance to smaller portfolio management without lacking quality of services.

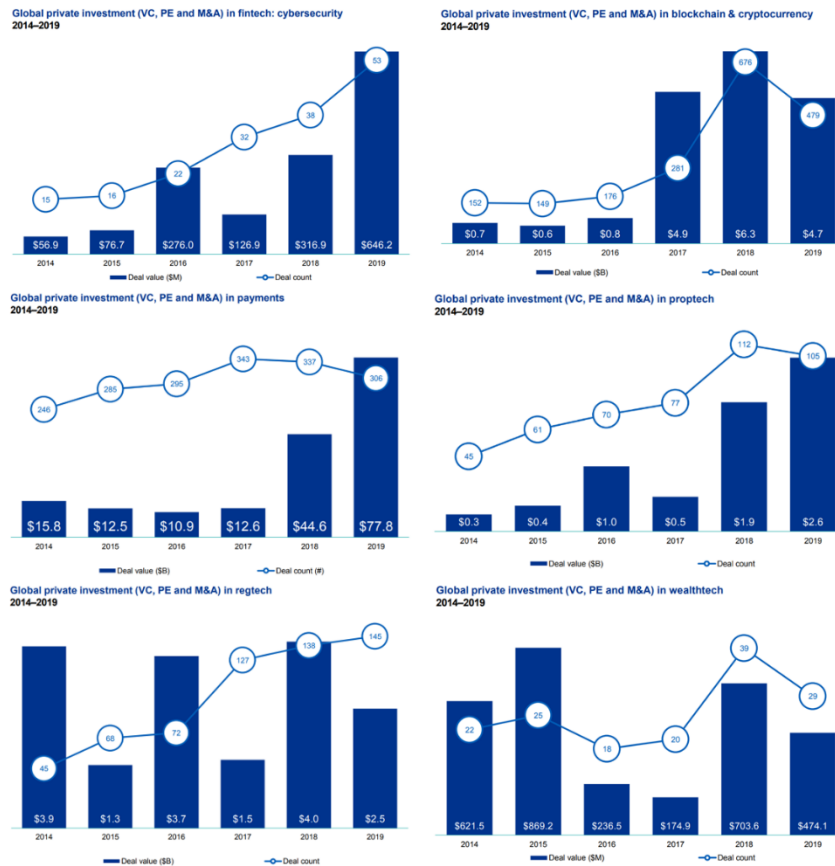


Figure 5 – Global yearly investment activity by different financial technology sectors. Source: KPMG.

Across all these sectors, there is a common denominator: the always growing use of AI. Historically, the Fintech industry has been among the earliest AI adopters. As of today, AI is becoming the main driver of digital transformation in traditional finance and the golden standard for FinTech services. In fact, according to a recently published report authorized by the World Economic Forum and conducted in partnership with the Cambridge Centre for Alternative Finance (Ryll, 2020)³³, by 2022, we can expect mass adoption of AI in the financial industry on a global scale. Otherwise, legacy financial services may become obsolete in as little as two years. AI and Data analytics go hand in hand, and technologies like Machine Learning, Neural Networks, and Natural Language Processing, continue to

32 <https://fintech-insights.com/mapping/wealthtech-europe/>

33 <https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/shifting-paradigms/#.YJUH8rUzZPY>

improve data-crunching capabilities for financial industry players³⁴. Today, finance is a business in cyberspace. Autonomous Agents (AA) and AI have already changed the landscape and will further change it in the next decade. Kensho’s Artificial Intelligence Investment Analyses platform has been acquired³⁵ by S&P for \$550M, the largest price tag on an AI engine. Both AA and AI are utilized in Augmented Finance where software-based reasoning is taking the lead. Although unnoticed by the general public, the digital transformation in the financial arena is changing the landscape.

As of 2018 in the US, there are some 2.5 million people working in financial institutions (including banking, investment, and insurance), and AA/AI is expected to have an impact in the range of \$1 trillion³⁶, \$490 billion of which in the front office, \$350 billion in the middle office and \$200 billion in the back office. Personal consulting agents, that can take the shape of Alexa, are now starting to revolutionize the front office.

As shown in Figure 6, the foreseen revenue out of AA/AI (in the US) may be reaching an aggressive estimate of \$500 B signalling a shift from the money saved using AA/AI to companies providing FinTech services through AA/AI (most of these companies are newly created companies). Notice that in the finance industry, the digital transformation is not acting on a product, rather on services and the processes used in the “manufacturing” of these services.

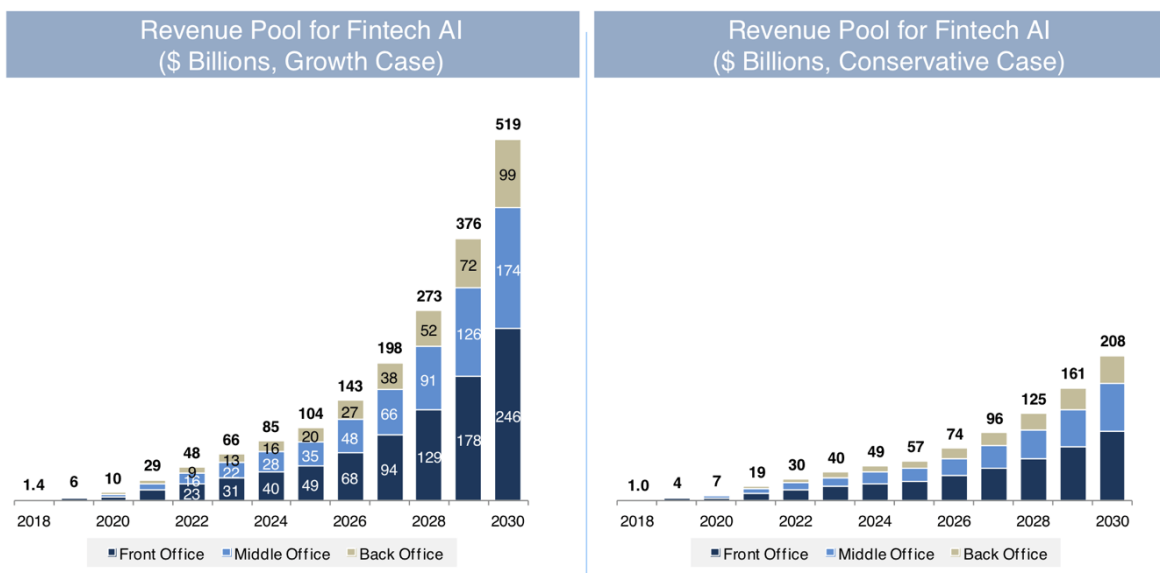


Figure 6 - Aggressive and conservative market estimates for autonomous systems and AI in the Fintech market – US market - Source: Autonomous Next

Although both AA and AI are the technology building blocks of Augmented Finance, AI takes the lion’s share since it is the one that is analysing data, evaluating risk and perspectives, and enabling more intelligent and automated processes, along with personalized services that are tailored to customers’ needs. This holds true for applications in different areas such as retail banking, corporate banking, payments, investment banking, capital markets, insurance services, financial services security. All of these applications

34 <https://coruzant.com/ai/how-ai-and-data-analytics-are-shaping-the-future-of-fintech/>

35 <https://next.autonomous.com/augmented-finance-machine-intelligence>

36 <https://autonomous.app.box.com/v/augmentedfinance>

leverage very large datasets from legacy banking systems (e.g., customer accounts, customer transactions, investment portfolio data), which they combine with other data sources such as financial markets data, regulatory datasets, real-time retail transactions and more.

AA are more executors, both locally and across a network. These characteristics of AA to exist locally and roam networks makes them applicable to a variety of contexts, becoming more a sort of commodity, hence limiting their market value. Besides, since they can be applied to a variety of contexts and they are network entities, it is foreseen that the major market value is found in telecommunications applications, particularly in the management of distributed resources. The global market value is estimated in \$345 M in 2019 to grow to almost \$3 B in 2024³⁷.

The Fintech transformation, as in several other areas hit by the Digital Transformation is not, per sé, creating new value, rather it is decreasing the overall market value. By disintermediating the front office, FinTech apps have saved \$5 B to US consumers which translates to a \$5 B decrease in the FinTech market.

Financial and related sectors are facing major competitive pressures, from increasing consumer demands to the emergence of new competitors. That results in more globalised and concentrated markets, although with more room for innovation and different value proposals. In addition, it is very likely that COVID-19 pandemic will have a deep impact in these sectors, boosting change and altering relationships with customers.

2.2 THE CHALLENGES OF DIGITAL TRANSFORMATION

The current **competitive landscape** of the financial sector as a whole has been shaped by a set of forces present for more than ten years, among them:

- **Diminishing profitability**, as a long-term trend aggravated by the Great Recession of 2008 from which financial and insurance companies never fully recovered;
- **Digitalisation**, both from the demand and from the supply side;
- **Globalisation**, expressed in terms on international market integration with common legislation and capital markets.

As an answer from the industries themselves, financial and insurance firms have been forced to an endless search of rationalisation, focusing on cost reduction and corporate concentration. At the same time, cross-sector competition has increased, with a steady increase of entrants from other industries - especially in the financial sector as it will be seen later - as well as start-ups and new players.

As for the **consumer perspective**, several trends are key to understand the evolution of attitudes and demands:

- **Higher service expectations**, as user experiences from other sectors and services (social networks, media and entertainment) based on technology are considered the standard for finance and insurance too;

³⁷https://www.marketsandmarkets.com/Market-Reports/autonomous-agents-market-201425821.html?gclid=CjwKCAjwscDpBRBnEiwAnQ0HQEg3oQjhgubL4mQ8Lctx5UHCxAZhXgw0l7-li9abYUHe2zNmyqX-FRoC0jQQAvD_BwE

- Growing awareness and activism **against the established market** power of financial institutions in services like payments or credits;
- **Collaborative and circular economy** applied to finance and banking services, with Peer-to-Peer (P2P) initiatives around blockchain and Lending Clubs.

These shifts can be partly explained by the rise of Millennials (people born between ca. 1980-1998) as a major consumer cohort with different consumer patterns than previous generations, Financial and insurance digital transformation takes advantage of Big Data, AI and IoT (*Internet of Things*) technologies in order to improve the accuracy and cost-effectiveness of their services, as well as the overall value that they provide to their customers. Nevertheless, despite early deployment instances, there are still many challenges that need to be overcome prior to leveraging the full potential of Big Data/IoT/AI in the finance and insurance sectors, which could also act a catalyst for attracting more investments and for significantly improving the competitiveness of enterprises in these sectors.

In particular, financial institutions and insurance organizations are currently faced with the following challenges:

- **Data Fragmentation and Interoperability Barriers:** Currently, most of the data collected and possessed by financial organizations reside in a wide array of fragmented systems and databases, including operational systems and On-Line Transactional Processing (OLTP) databases, On-line Analytical Processing (OLAP) databases and data warehouses, data lakes (e.g., Hadoop-based systems) and others. In this fragmented landscape, heavy analytical queries are usually performed over OLAP systems, which lead financial organizations to transfer data from OLTP, data lakes and other systems to OLAP systems based on intrusive and expensive Extract-Transform-Load (ETL) processes. In several cases, ETLs consume 75%-80% of the budget allocated to data analytics while being a setup to seamless interoperability across different data systems using up-to-date data.
- **Regulatory Barriers:** Big Data and IoT deployments must respect a complex and volatile regulatory environment. In particular, they must adhere to a range of complex regulations, e.g., PSD2, MiFIDII/MiFIDR (Markets in Financial Instruments Directive), 4AMLD/4MLD (4th EU (Anti) Money Laundering Directive) for financial/insurance, while at the same time complying with general regulations such as the GDPR and the ePrivacy directive. To this end, there are a number of initiatives that aim to establish regulatory sandboxes³⁸ i.e., specialized environments that facilitate experimentation through ensuring processing of data in-line with applicable regulations. Nevertheless, such sandboxes are in their infancy, demand the engagement of regulators and are not tailored to leading edge Big Data/IoT applications.
- **Data Availability Barriers:** In order to innovate with IoT and Big Data, financial and insurance organizations (including FinTech/InsuranceTech innovators) need access to realistic datasets (e.g., customer account and payments' datasets) that would allow them to test, validate and benchmark data analytics algorithms. Such data are hardly available, as their creation requires complex anonymization processes or even tedious processes for simulating/synthesizing them. Therefore, innovators have no easy way to access data for experimentation and testing of

38 [FCA15], [Sandboxes18]

novel ideas. Moreover, the fragmentation of European FinTech/InsuranceTech ecosystem is a challenge for sharing such resources across financial/insurance organizations and innovators;

- **No Validated Business Models:** Big Data/IoT deployments in finance/insurance have in several cases demonstrated their merits on the accuracy, performance and quality of the resulting services (e.g., increased automation, improved risk assessment, faster transaction completion, better user experience). However, there is still a lack of validated business models that could drive monetization and deliver tangible business benefits.
- **Impact of new technologies on banking business models:** (example of insurance and IoT) considering this, it's also important to consider the practices about risk assessment models that may be impacted from new technologies. In this context banks and insurance may consider how to reshape business enabling positive value for banks and the customer.

2.3 DIGITAL TRANSFORMATION AREAS IN THE FINANCIAL SECTOR

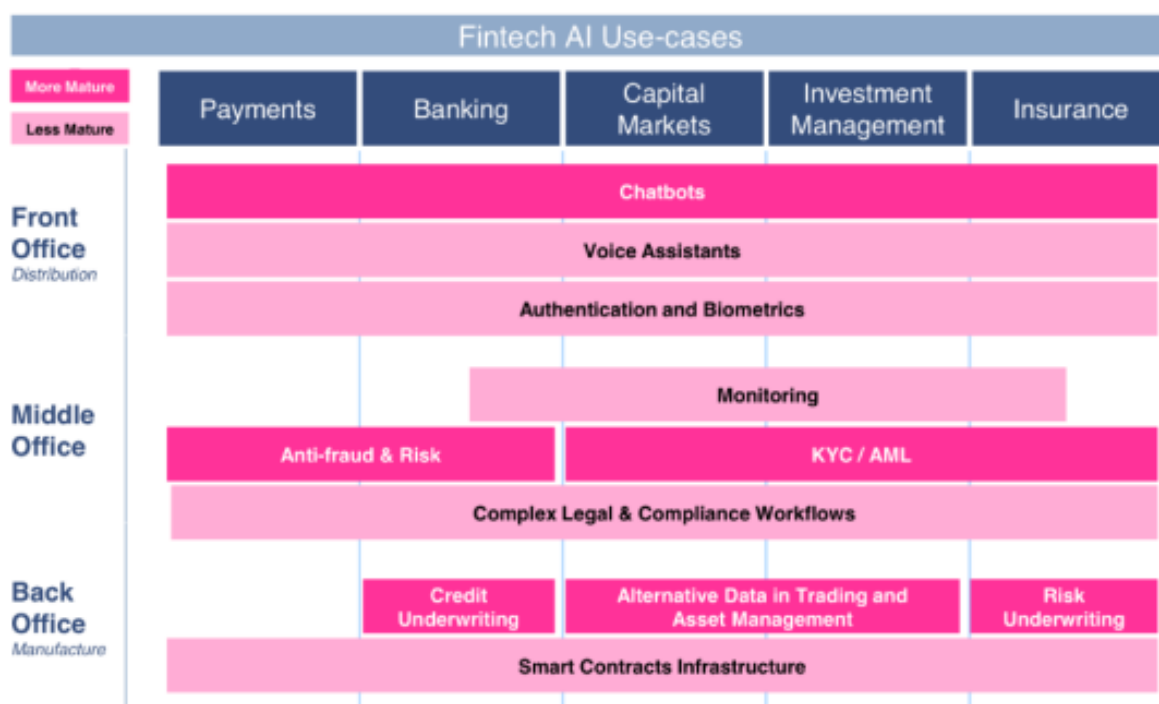


Figure 7 - Various segments of Autonomous Agents and Artificial Intelligence use in Fintech.

Source: Autonomous Next

The vast majority of digital transformation applications for the finance and insurance sectors are data-intensive. This holds for applications in different areas, such as retail banking, corporate banking, payments, investment banking, capital markets, insurance services, and financial services security.

All of these applications leverage very large datasets from legacy banking systems (e.g., customer accounts, customer transactions, investment portfolio data), which they combine with other data sources such as financial markets data, regulatory datasets, real-time retail transactions and more.

With the advent of IoT devices and applications (e.g., fitbits, smart phones, smart home devices), several FinTech/InsuranceTech applications can take advantage of contextual data in order to offer better quality of service at a more competitive cost (e.g., personalized healthcare insurance based on medical devices and improved car insurance based on connected car sensors).

Furthermore, alternative data sources (e.g., social media and online news) provide opportunities for new more automated, personalized and accurate services. Moreover, recent advances in data storage and processing technologies including advances in AI and blockchain provide new opportunities for exploiting the above-listed massive datasets and are expected to stimulate more investment in digital finance/insurance services.

2.3.1 Digital Finance State of the Art Solutions

Consumer pressure seems to be a major driving force in banking services, with mounting levels of demands by consumers³⁹. Traditional banks lead the way in sophisticated segments, such as International Trade or Investment Banking, whereas challenger banks seem to have raised the bar in simple services.

The evolution of corporate operations is a result of the observed increased importance of Fintech, in European Union as in the rest of the world. **Figure 8** shows that the total value of financial operations reached a maximum in 2019, while at the same time the number of transactions declined slightly, which seems to outline some consolidation of the sector (less deals but of a higher value).



Figure 8 - Corporate operations in Fintech, 2014-2019, source: KPMG, Pulse of Fintech H2 2019

Additionally, players from different sectors have been acquiring an increasingly roles, particularly:

- BigTech, or Large Technological companies, such as the so-called “GAFA” (Google, Amazon, Facebook and Apple), as well as Alibaba. They tend to position

³⁹ Cap Gemini and EFMA, World FinTech Report 2020

as convenience, low added value suppliers, leveraging on synergies with their current activities and consumer bases;

- Telecommunication firms, with specific technology-intensive strengths (IoT, mobile payments) which allow a different approach to banking services.

In practice, the degree of adoption of non-traditional banking solutions in each country depends on several factors, such as the degree of ubiquity of established banks or the market power of telecommunication companies. Thus, the Chinese market boasts the highest volume of Fintech credit per person (\$324, three times more than in the USA or UK); while in less developed countries like Argentine or Brazil Telecommunication companies stand for more than 40% of total Fintech assets, as opposed to less than 5% in the American or British markets⁴⁰.

Big Data, AI and IoT enable a shift of the financial/insurance sector towards products and services, which will significantly improve the competitiveness of the sector and will stimulate more investment in Big Data and IoT driven innovation. Disruptive solutions based on AI and related technologies are arising in different contexts of financial services; the following section presents an overview of innovative solutions that will impact the digital finance sector in the near future.

Risk Assessment

Credit Risk Assessment for small and medium enterprises (SMEs): most banks consider SMEs high-risk customers, which limits their ability to lend even to the bigger and wealthier SMEs. This is largely a result of Basel III ⁴¹and IFRS9⁴² requirements but also a consequence of the fact that SMEs: (i) Have very limited coverage by credit reporting service providers; (ii) Are subject to weak contract or bankruptcy laws and judiciaries; and (iii) Are characterized by high informality in developing markets. This limitation has an adverse socio-economic, as SMEs contribute to more than 50% of the EU economies. Banks are in need of a novel approach for accessing SMEs' credit risk, such as the sharing of large volumes of data across financial organizations and their processing based on AI techniques. Innovative solutions will integrate AI and blockchain-based systems for credit risk scoring of SMEs, enabling the collection and sharing of both traditional data (i.e., banking, accounting, transactional, and sales data) and various forms of alternative data (i.e., online ranking, news and blog feeds and social media, mobile, and individual data).

Credit Risk Assessment for Investment Banking: In today's lower return / higher risk business environment, one of the main challenges in asset management is to provide detailed risk information in a timely manner, i.e., in real time. Financial organizations and asset managers are therefore seeking novel ways for overcoming the current practice of risk applications relying on batch processing to produce aggregated reports. Alternatively, traders, risk managers, and sales negotiators should be supported on the fly, which requires real-time analytics technologies to process live operational data. Innovative real-time risk assessment and monitoring solutions will leverage SAS (Statistical Analysis System) technologies for measuring various types of risk, and above all, market risk of portfolios of assets.

40 The Economist, 21-Nov-2019, "Plug and play: Big Tech takes aim at the low-profit retail banking industry"

41 <https://www.bis.org/bcbs/basel3.htm>

42 <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-9-financial-instruments/>

Personalized Retail and Investment Banking Services

Financial institutions currently have access to very large amounts of customer-related data from many different data sources, including both banking systems and alternative data sources (such as open data and social media). SAS technologies give the opportunity to aggregate, consolidate and share such data across institutions/organizations. Financial organizations are offering opportunities that are increasing the automation, accuracy and credibility of customer-centric processes, including KYC/KYB (Know Your Customer/Know Your Business), services personalization, credit risk scoring and more.

Financial Crime and Fraud Detection

Financial crime costs the economy billions and facilitates drug trafficking, slavery and prostitution. An estimated £57 B is laundered through London per year. In 2016, at any given time, an estimated 40.3 M people worldwide were in modern slavery, including 24.9 M in forced labour and 15.4 M people in forced marriage. The UK Government's organized crime strategy indicates that drug trafficking costs the UK an estimated £10.7 B per year. In addition to the extreme cost to society, financial crime costs the business in terms of reputation and fines. Over the last few years, financial institutions have been fined billions. Using advanced computing techniques could start to change this environment by creating a more up-to-date, accurate, versatile, and complete view of the customer's profile and behaviour. This level of information could reveal abnormal behaviour earlier and more accurately and enable the business to intercept criminal behaviour more effectively. In the financial crime intelligence scene, Machine Learning (ML) has the ground-breaking potential to reveal much more realistic financial crime typologies, compared with traditional rule-based systems.

Insurtech: Personalized Usage Based Products

Insurance has been considered as a technologically conservative industry. However, new technologies are already transforming the concept of insurance itself, since they are changing the way to detect or even prevent risks, as well as to measure impacts⁴³. The use of data inference and ML methodologies are already leading to proactive insurance approaches and to new services which blend insurance with risk management⁴⁴.

At the same time, customer relationships in insurance are easier to virtualize than in banking as processes are more streamlined with less touchpoints. Therefore, the use of online channels for product comparison and purchasing insurance policies is widespread, according to a worldwide survey conducted by Cap Gemini and EFMA⁴⁵.

In any case, the financial impact in terms of corporate operations related to Insurtech in the world is much smaller than in Fintech, though the trend is clearly positive, as shown in

Figure 9:

⁴³ The Economist, 29-Mar-2019, "Counsel of protection - The coming revolution in insurance"

⁴⁴ HBR, 29-Apr-2020, "The Case for AI Insurance"

⁴⁵ Cap Gemini and EFMA, World Insurance Report 2020



Figure 9 – “Insurtech private operations, 2014-2019” - Source: KPMG, Pulse of Fintech H2 2019

The convergence of vehicle-derived IoT and Big Data stemming from a massive number of drivers from all around the world is already revolutionizing transport services and holds the promise to disrupt the ways relevant insurance products are personalized and delivered to customers. However, most attempts towards offering pay as you drive programs rely typically on the installation of relatively simple OBD2 dongles, which are only able to collect data from a very limited number of vehicle signals. They assess the driver behaviour analysing acceleration, steering drive, speed, and brake patterns, resulting in a coarse-grained classification. A large-scale system relying on high-quality vehicle batch and streaming data could allow for consideration of many more factors that suppose a fundamental factor at the time of establishing the risk of a drivers, e.g., quality of the roads in his geographical area (road roughness), identification of geographically aggregated driving patterns or the status and maintenance of the car. There may also be benefits for other models like usage-based insurance or even to increase the capacity for fraud detection. Moreover, the collection of additional data from connected vehicles (e.g., city pollution data from air-quality sensors of the cooling systems) can provide even more opportunities for added value insurance products and services.

2.4 THE IMPACT OF THE COVID-19 PANDEMIC ON THE FINANCIAL SECTOR

From the beginning of the global pandemic, a set of forces have arisen at different levels, influencing technology, finance and insurance. It is impossible to foresee yet how deep will be the impact of these forces, as eventual changes will depend on:

- Size and evolution over time of each force
- Interaction among them
- Specific level of disruption on Finance and Insurance

For the case of consumers, some interesting insights can be obtained from ongoing research on individuals and SMEs in the Spanish market⁴⁶:

⁴⁶ Inmark barometer- Impact of COVID-19 in Spain (Apr-20 – to date)

- During the crisis caused by the COVID-19 outbreak, technology is not perceived as a key issue, but as an enabler to face health or financial problems;
- Most individuals and SMEs do not perceive financial institutions as key partners for dealing with the consequences of the COVID-19 pandemic, with rather low levels of awareness about support initiatives or new services;
- In the case of Individuals, the results suggest an increased search for easier ways to consume, to pay, to finance and to insure;
- As for SMEs, the most important concern is to accelerate their process of transformation themselves, with an extra effort on e-commerce and the transactional and finance needs associated to it.

3 EVOLVING DIGITAL FINANCE TOWARDS 2025

The Digital Finance world is evolving at fast pace by following the continuous improvements in technologies to apply innovative solutions addressing the market needs. However, many of the monolithic infrastructures associated with the traditional “Big” Financial Institutions are built upon regulations and structured processes that appear not completely to fit with the digital transformation, enhancing the need for resilient and flexible infrastructures. Legal and ethical compliance are pillar for the Financial Sector and addressing such challenges may be facilitated by innovative framework such as the regulatory sandboxes.

The future of Digital Finance is data-driven and relies on fast and reliable solutions, paving the ground for optimization and intelligent automation of its models. For example, AI technologies aim to lighten and improve such processes by bringing innovation and facilitating access and safe deployment of solutions, although facing high barriers for its adoption. Indeed, accelerating such implementation is key for success: innovative hubs and the establishment of resilient infrastructures could indeed speed up processes. Building a resilient new era means leveraging on and exploiting the enablers brought by the innovation facilitators such as Open and Digital Innovation Hubs.

On the one hand, applications need to be tested in a safe and secure environment, thus testbeds and sandboxes could facilitate the deployment by relying on cloud infrastructures and reducing the risks. On the other hand, real life scenarios can bring the know-how and allow a faster adoption of new technologies and solutions. Therefore, the main objective of this paper is to share and enhance the impact of the financial sector highlighting the current innovation, the challenges, and the vision towards 2025.

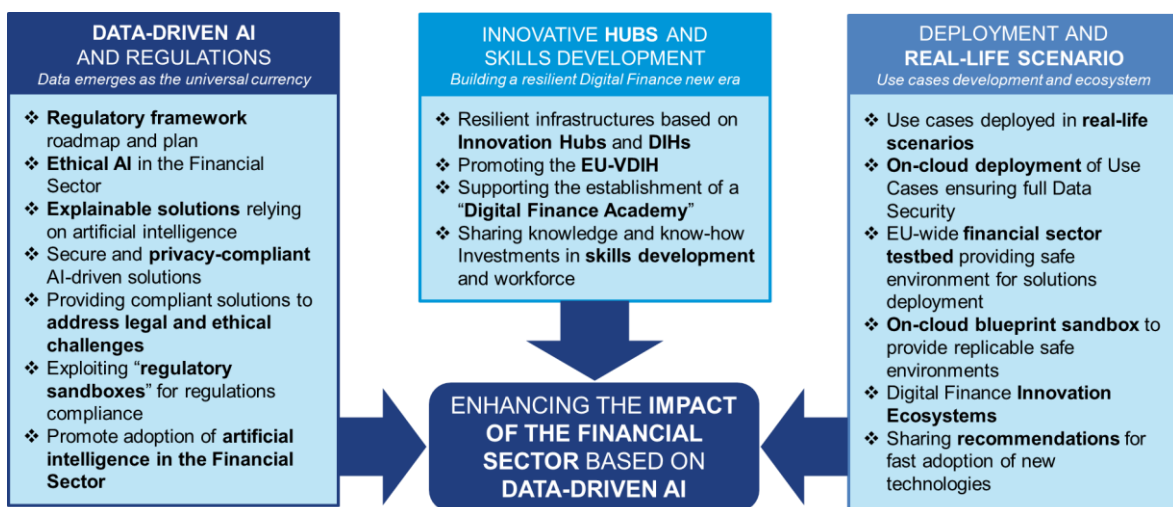


Figure 10 - Impact of the data-driven AI in the financial sector

In Annex I “Deep Dives on projects & applications” interesting examples can be retrieved, including:

- the increasing number of new digital native banking entities (the so-called “Digital Native Banks”, where banking services are offered mainly through the digital channel, supported by a lean technological architecture (without the burden of legacy systems) exploiting the latest innovations on data management, and thereby offering an innovative user experience (more details in I.1);

- the Decentralized Finance (DeFi), that regards the improvement of an ecosystem of financial services, developed on top of distributed infrastructures (such as blockchain), in order to grant a transparent and trustless framework, and capable of addressing some of the critical aspects of the traditional financial paradigm, supporting innovation (more details in I.2);
- the application to data governance and management in AI and ML of the decentralized approaches, based on consensus, smart contracts and distributed ledger technologies (more details in I.3);
- the use of network analysis relating to the study of financial relations. Network analysis is a specific branch of network science focusing on the study of complex networks through a wide range of tools and methodologies and representing a realm that is not easy to define or to position. Its application to financial relations gives rise to a systemic and network approach designed to ease liquidation of trade receivables, optimise short-term credit allocation, and enhance and improve the efficiency of risk analysis models (more details in I.4).

3.1 RESPONSIBLE DATA-DRIVEN AI

The section mainly deals with Data and its limitations, problems, and challenges to feed reliable AI algorithms.

High relevance will be given to innovative data models, data spaces, different approaches to exploit the use of real datasets (anonymization, synthetic data, etc) and the trustworthiness and explainability of the algorithms, as well as their bias, based on different data sources.

The value of an Open-Library of ML algorithms, such as the one that Infnitech will produce, will also be described.

Finally, ethical challenges related to the interactions among humans and machines are addressed as well, in their complementary role in respect to regulations in driving the possible development of AI and the utilization of Data in the financial domain.

3.1.1 Data for AI

The attention to AI technology in the finance industry is quite profound. Over the years the combination of AI and Big Data started to reveal its possibilities allowing to transform raw data into actionable business intelligence.

Together with the algorithms, Data is a fundamental element for the training of an AI system and a crucial aspect for its performance evaluation with a view to enhancing its efficiency. However, resulting insights from poor Data quality cannot be trusted. It is therefore imperative the implementation of a structured approach to AI with a proper information architecture.

The financial sector, from this point of view, boasts a long history. Since the beginning of the last decade, financial institutions/banks have embarked on a long process to develop data governance practices. In the early years, the primary need was to implement tools, processes and techniques capable to produce an adequate quality to meet regulatory requirements in specific areas.

Gradually, in the following five years, a more encompassing and multifaceted vision of the subject took hold, moving from Data Quality to the concept of Data Governance, instrumental to ensure a consistent and mature supervision of company information. Also

in this case, a strong push came from the regulatory evolution. It should be noted that, in this phase, the efforts made to adapt to the regulatory requirements provided a valuable opportunity to grow, as they drove the management focus on the need to govern data and manage company information.

Nowadays, information represents a capital of great value for the financial institutions and other industries alike. They represent an incentive to maintain a competitive advantage.

Though very close, Data Governance and AI/ML are two separate disciplines, where Data Governance has the task to provide quality data to the entire company, including for the use of AI solutions.

Typically, there are four macro-areas that characterize the Information Governance corporate standards.

The first important element concerns the formalization of the principles and guidelines that need to be taken into consideration when setting up the model. One of the most common approach consists in structuring a set of company policies and standards describing the fundamental pillars of the Information Governance model, and the definition of a roadmap related to the implementation of the various projects in the Information Governance field.

A further relevant element concerns the methodologies that support the various information governance activities. In this context, an important aspect is the measurement, which is typically applied according to a double perspective: to assess the state of maturity and evolution of the various projects, and to monitoring the individual areas of implementation of the overall Information Governance model. To date, attention to measurement focuses mainly on Information Quality metrics.

With regard to the definition of the organizational structure and the roles involved, there are very different approaches, according to the specific characteristics of each reality.

However, there is a large tendency to identify some macro-types of actors who play a key role in Information Governance processes: an Information Governance unit, which has the main task of supervising and coordinating the activities defined in the company policies; a business contact person, who assumes responsibility for a specific information perimeter considering its business purposes; an IT contact person, who oversees the IT applications and procedures. Other typical roles that can be identified in Information Governance systems are for example: the Data Steward, the IT Security Contact and the Contact Person for outsourced IT activities. Due to the intrinsic characteristics of the information itself, which comes from different sources and is used by different users, it is important to evaluate the methods and purposes with which the various corporate organizational entities access, modify and use the information. For this purpose, it is particularly necessary to identify the most suitable organizational location of the Information Governance unit. This choice typically derives from company objectives and priorities, weighted on the basis of additional factors such as the level of maturity of the organization or the IT sourcing model adopted.

Ultimately, it is important to consider the profile of the tools used in the Information Governance system. Next to the choice between the different technologies to be adopted, it is necessary to pay due attention to the underlying architectural model and to the integration dynamics of the different tools adopted, in order to have and synergistically use the potential of each of them and to achieve the best performance

3.1.2 Security and Privacy

Most applications of AI require huge volumes of data in order to learn and make intelligent decisions. AI is high on the agenda in financial services due to its potential for radically

improved services, commercial breakthroughs and financial gains. This will and is to be done within a framework of increased regulatory burdens requiring better governance, transparency and security around data and the services that are being provided.

The further extension of regulations like GDPR are furthermore designed to preserve individuals' fundamental privacy rights which is having a transformational impact on how data is collected, retained, shared, and processed.

A baseline of security requirements is critical in order to engender trust from customers and to progress with the needs of the various regulatory and legal obligations incumbent on a financial institution. These requirements typically need to fulfil the need to maintain confidentiality, availability and integrity within the systems, applications and services being provided.

The adoption of security-by-design is essential for maintaining the confidentiality, integrity, availability, and resilience of the data held by the organization. The adoption of privacy-by-design is key to demonstrating appropriate processing of personal information and reducing the risk of data breaches.

Any use of AI within the wider context of digital finance whether it be for automated decision making or customer profiling should ensure that the collection, storage, transmission, and processing of data meet security- and privacy-by-design principles to prevent data leakage, whether by accidental or unlawful destruction, loss, alteration, unauthorized disclosure, intrusion, accidental or unauthorized data exposure. The adoption of security- and privacy-by-design is an essential principle to minimize risk of data loss, breach, harm to individuals or other infractions.

Security needs to be incorporated throughout the entire lifecycle. As new AI applications and use cases emerge, devices and infrastructure that run these applications need to be capable of adapting to an evolving threat landscape. To address the high-grade protection requirements, security needs to be multi-faceted and "baked-in" from the edge devices incorporating neural network processing system-on-chips (SoCs), right through to the applications that run on them and carry their data to the cloud.

At the outset, system designers adding security to their AI product must consider a few security enablers that are foundational functions that belong in the vast majority of AI products, to protect all phases of operation: offline, during power up, and at runtime, including during communication with other devices or the cloud. Establishing the integrity of the system is essential to creating trust that the system is behaving as intended.

Secure bootstrap establishes that the software or firmware of the product is intact ("has integrity"). Secure bootstrap systems use cryptographic signatures on the firmware to determine their authenticity. Flexibility for secure boot schemes is maximized by using public key signing algorithms with a chain of trust traceable to the firmware provider. Public key signing algorithms can allow the code signing authority to be replaced by revoking and reissuing the signing keys if the keys are ever compromised. The essential feature that security hinges on is that the root public key is protected by the secure bootstrap system and cannot be altered. Protecting the public key in hardware ensures that the root of trust identity can be established and can't be compromised.

Encryption algorithms can still be compromised if the keys are not protected with key management. For high-grade protection, the secret key material should reside inside a hardware root of trust. Permissions and policies in the hardware root of trust enforce that application layer clients can manage the keys only indirectly through well-defined APIs.

Continued protection of the secret keys relies on authenticated importing of keys and wrapping any exported keys.

Whether in the cloud or at the edge, AI applications will continue to get more sophisticated, and data and models will need to be updated continuously, in real time. The process of distributing new models securely needs to be protected with end-to-end security. Hence it is essential that products can be updated in a trusted way to fix bugs, close vulnerabilities, and evolve product functionality. A flexible secure update function can even be used to allow post-consumer enablement of optional features of hardware or firmware.

Big Data can often have significant privacy concerns associated with them so protecting large data in memory, such as DRAM memory, or stored locally on disk or flash memory systems, are essential. High bandwidth memory encryption (usually Advanced Encryption Standard – AES – based) backed by strong key management solutions is required. Similarly, models can be protected through encryption and authentication, backed by strong key management systems enabled by hardware root of trust.

To ensure that communications between edge devices and the cloud are secured and authentic, designers use protocols that incorporate mutual identification and authentication, for example client authenticated Transport Layer Protocol (TLS). The TLS session handshake performs identification and authentication, and if successful the result is a mutually agreed shared session key to allow secure, authenticated data communication between systems. A hardware root of trust can ensure the security of credentials used to complete identification and authentication, as well as the confidentiality and authenticity of the data itself.

The principle of data minimization requires that the data used shall be adequate, relevant and limited to what is necessary for achieving the purpose for which the data is processed. This means that a controller cannot use more personal data than necessary, and that the information selected must be relevant to the purpose. A challenge when developing AI is that it may be difficult to define the purpose of processing because it is not possible to predict what the algorithm will learn. The purpose may also be changed as the machine learns and develops. This challenges the data minimization principle as it is difficult to define which data is necessary.

However, data minimisation is more than a principle limiting the amount of detail included in training or in the use of a model. The principle also stipulates proportionality, which restricts the extent of the intervention in a data subject's privacy that the use of personal data can involve. This may be achieved by making it difficult to identify the individuals contained in the basic data. The degree of identification is restricted by both the amount and the nature of the information used, as some details reveal more about a person as others. The use of pseudonymisation or encryption techniques protects the data subject's identity and help limit the extent of intervention. This principle also forces developers to thoroughly examine the intended area of application of the model to facilitate selection of relevant data necessary for the purpose. Furthermore, the developer must consider how to achieve the objective in a way that is least invasive for the data subjects. The assessments performed need to be documented, so that they can be presented to the Data Protection Authority in the event of an inspection, or in connection with a preliminary discussion.

3.1.3 The evolving Regulatory Framework fostering EU competitiveness

The EC, in accordance with its 2018 FinTech Action Plan and with the aim of favouring the scaling-up of the Digital Finance in the EU, appointed the **Expert Group on Regulatory Obstacles to Financial Innovation (ROFIEG)**. It was responsible for reviewing the European legal and regulatory framework applicable to the financial sector and for analysing the extent to which such framework is technology-neutral and capable of strengthening the innovation in the domain or if adaptation would be advisable. ROFIEG's activity and survey were also directed to ensure financial stability integrity, as well as consumer and investor protection in light of the new opportunities and risks brought by the use of AI solutions, Big Data Technologies and DLT/Blockchain in the financial sector in several contexts, such as for instance, client interfaces or other applications (credit scoring, chatbots, robo-advice, insurance pricing and underwriting).

Though such technological innovations are common to different areas of economic and social life and across the public sector, the regulatory framework for the Financial Services might imply specific challenges and require tailored responses.

The ROFIEG published in December 2019 “**30 Recommendations on regulation, innovation and finance. Final report to the European Commission**”. It comprises insights in light of the need to ensure the global competitiveness and sovereignty of the EU in a global technology-enabled financial market. ROFIEG's efforts were also directed to fill the big gap existing between the EU on the one hand (5%) and the US (65%) and the PR China (35%) on the other hand, in relation to the current distribution of the value of technology companies. The underlying vision is that fostering a proactive lead of the EU in this sector would allow to promote its fundamental European values, such as data privacy and fair competition, particularly important considering the current less stringent standards of operation of fast developing non-EU financial markets in areas of law such as data protection. In fact, the ROFIEG acknowledges that, if appropriately calibrated, in the long run these areas of law allow to protect ethical “values whilst not posing an undue barrier to the innovative use of technologies”. This would be facilitated especially if the EU is capable of providing genuine thought leadership for the regulation of innovative technology in the financial sector, setting the basis for EU competitiveness and regulatory sovereignty in relation to technology-driven finance.

The **recommendations and findings** identified by the ROFIEG include, among others:

- the need to adapt the legislation to respond to new and changed risks caused by the use of innovative technologies, such as AI and DLT, and, in general, the **need of a harmonised framework in the EU**, taking up any emerging opportunities with respect to RegTech or SupTech (Recommendations 1-12);
- the need to **remove regulatory fragmentation** and to ensure a **level playing field** between the different market players across the entire EU, including FinTech start-ups and BigTech firms (Recommendations 13-24);
- the necessity to reconcile the **regulatory regime of personal and non-personal data**, including data sharing, with the risks and opportunities brought by the financial innovation (Recommendation 25-28);
- the need to foster **financial inclusion** and the **ethical use of data** (Recommendations 1 and 29-30).

Besides stressing the relevance of international cooperation in setting adequate standards, the ROFIEG identifies the following **priorities in terms of regulatory reform** for establishing a level playing field amongst actors and promoting innovation:

- **explainability and interpretability** of technology, in particular AI, functional to protect consumers and businesses and facilitate supervision (Recommendation 1);
- the regulatory recognition that activities creating the same risks should be governed by the same rules (Recommendation 13);
- the needs to stop and **avoid regulatory fragmentation**, especially in the area of customer due diligence (CDD) / know your customer (KYC) (Recommendations 15-17);
- maintain **consumer choices, fostering innovation and preventing unfair treatment** of competing downstream services by large, vertically integrated platforms (Recommendation 22);
- a strengthened **framework for access to, processing and sharing of data**, in order to promote innovation and competition (Recommendations 27 and 28).

In line with these recommendations and building on the work carried out in the context of the 2018 FinTech Action Plan and the work of the European Parliament, European Supervisory Authorities (ESAs) and other experts, the EC intervened in the financial sector promoting an organic framework of the matter, with the adoption of the **Digital Finance Package** (DFP) already mentioned in this White Paper, with the double ambition of driving digital finance with strong European market players in the lead and in the meantime of upholding European values.

The DFP includes the programmatic lines for digital finance development and a series of legislative proposals on crypto-assets and digital resilience. As underlined by the European Banking Federation, “the DFP is a major step towards the comprehensive regulatory framework for financial services that will help make the European Union fit for the digital age”.

In line with ROFIEG’s recommendations, the DFP was adopted in consideration of the need for a cohesive, technologically-neutral and innovation-friendly EU financial services framework. One of its main aims is to ensure that the European regulatory framework can facilitate innovation in the interest of consumers and market efficiency, moving towards the creation of a **European financial data space** by addressing the new challenges and legal risks to digital transformation and strengthening technological operational resilience, in the financial domain. The DFP highlights the importance of safeguarding financial stability and of applying the principle “same business, same risk, same rules”, as well as of promoting operational resilience of the financial ecosystem with common minimum-security requirements, governed by a risk-based approach.

The harmonization of onboarding regulations, the interoperability of digital identities, and the strengthening anti-money laundering legislation with regard to customer due diligence were considered by the EC as pivotal to achieve digital market harmonisation. In addition, the pillars of the **Commission strategy by 2024** include the implementation of the so-called principle of the passport and one-stop shop system, with free access of citizens and businesses to cross-border services offered by companies supervised by other Member States, as well as a comprehensive regime supporting the use of DLT and crypto-assets in the financial sector. The EC also stresses the importance of the cloud computing infrastructure and the promotion of investments in software and AI applications. Likewise, it is key to have in place an adequate regulatory landscape for financial services based on AI, whilst adapting the legislation to the principle of technological neutrality. On the other hand, the trajectory is moving for the promotion of the data sharing between companies in the financial sector by implementing, by 2024, an open finance regulatory framework aligned with the European data strategy and other regulatory sources like the EU Payment

Services Directive. Another priority is facilitating digital access to financial data in real time providing, by 2024, that financial information is made public and communicated in standardised and machine-readable formats. An additional hint is that innovative technologies should include not only RegTech (technologies applied to regulation) but also SupTech (technology applied to supervision).

3.1.4 Ethical AI in the Financial Sector

The rapidly evolving financial landscape is characterized by the increasing and pervasive use and impact of autonomous and intelligent systems: the overall transformation underpinning this sector is characterized by high value AI applications, systems and intelligent workflows which use client / member data for generating new financial products and services delivering value. Such applications and systems are increasingly allowed to make relevant decisions without putting humans in the centre and this might hinder the advancement and uptake of AI and Autonomous Systems (AIS).

In fact, the mentioned transformation offers great promises and opportunities but at the same implies the risk of great emerging perils and biases, such as risks related to fairness, explainability, and privacy in data and algorithms.

The use of AI has launched vivacious debate, also raising ethical questions. If certain objections are understandable, it is nonetheless crucial to contextualise certain issues, remembering that the concept of ethics varies widely among individuals, societies, and jurisdictions, and can change over time.

Thus, it is important **to use ethical considerations as general recommendations**, which are abstract and flexible, aimed at favouring a principles-based approach, rather than as rigorous rules that risk suffocating innovation or becoming obsolete as culture and technology evolves.

It is just as important that politicians in charge and society in general take a **neutral position** on this technology, viewing its application, intentions, and underlying purposes objectively. This gives rise to the importance of promoting thought and discussion on a global ethical framework for AI, in order to ensure consumers and individuals' trust in this technology.

Lastly, it is worth noting the importance of following an **ethical approach in planning the development of the new systems**. This means keeping in mind, from the start of the phase of planning a new application, what ethical standards must be followed and what the ethical benchmarks will be.

Another factor that is often debated regards the need that the application of AI tools and technologies guarantee **fairness**. In general, there is no standard definition of fairness for the decision-making process, as a judgment involves ethical, legal, political, social, historical and cultural considerations and most of the time entails a decision that is a compromise.

Thus, dealing with the issue of fairness in AI means ensuring that there is a fair distribution of costs/benefits while guaranteeing that individuals and groups are free from prejudice, discrimination, and unjust stigmatisation. Moreover, the use of AI systems should never mislead or damage the free choices of individuals. To that end, a branch of research into AI systems has been created to investigate how to identify any prejudices in datasets and the ethical dilemmas posed by this technological development.

It is therefore critical to tackle with such risks and perils in an ethically-sound manner, in order to ensure transparency, minimize biases, encourage accountability. The upholding and operationalization of European values through practical and scalable approaches

aligned with customer expectations, ethical mandates and regulatory requirements, is expected to contribute to build trust and AI uptake in the financial domain: the responsible and ethically-driven design, development and adoption of AI and Big Data solutions is an enabling factor for accelerating the wider AI system deployment and large adoption in the financial domain, in alignment with the expectations of clients, employees, and society at large.

In this spirit, the IEEE has recently published in 2021 a high-value document, the “**IEEE Finance Playbook – Version 1.0. Trusted Data and Artificial Intelligence Systems (AIS) for Financial Services**”, which was elaborated in the framework of a dedicated initiative inside IEEE⁴⁷ to serve as guidance and specific and actionable tool for this area.

The IEEE consulted with several data and AI ethics experts in the financial services industry for developing the Playbook, so that to provide insights on the key ethical themes and findings relevant for the financial sector, accompanied by actionable recommendations.

From a wider perspective, the Playbook is related to the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, which conceived the “Ethically aligned design” concept. Such concept was generated in an ecosystem of more than 700 global experts who elaborated the “Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems”. It is a methodology leveraging the ethical and values-based design principles into design, development, and implementation of autonomous and intelligent systems, relying on several general principles including human rights, well-being, data agency, effectiveness, transparency, accountability, awareness of misuses, which should guide in the AI design and development pathway.

The Playbook represents a key step forward in the direction of the materialization of the ethical principles and requirements in this sector, providing a comprehensive framework and sector-specific guidelines built on top of strong ethical foundations, capable of enabling organizations and institutions operating in this sector to become more confident and proactive actors for driving economic and social value. The Playbook is also strongly aligned with European vision. In fact, for instance, it makes several references in most of its sections to the EU final Assessment List for Trustworthy Artificial Intelligence (ALTAI) and related ethical principles, which have many commonalities with the Ethically Aligned Design principles.

Therefore, after careful, in-depth examination of the content and insights of the **Playbook**, we believe that it represents a **comprehensive, noteworthy cornerstone for the ethical discussion** in the financial sector, to be considered as a starting point for the future advancement of the subject, where a relevant amount of work is still to be done beyond the current narrative and global discourse regarding ethics implementation in this sector, led predominately by Big Tech (e.g., IBM, Google, Microsoft). Therefore, in this paragraph we are taking excerpts from the Playbook for providing an overview of the most valuable findings in relation to the ethical discourse in the financial sector.

As regards **ALTAI (Assessment List for Trustworthy Artificial Intelligence)**⁴⁸ **list**, which was presented by the AI HLEG (Artificial Intelligence High-Level Expert Group On Artificial Intelligence) and provides an initial approach for the evaluation of trustworthy AI building on the Ethics Guidelines for Trustworthy AI, it can be used by financial organizations and institutions for self-assessment in their design and implementation of AI applications in order to take into consideration the peculiarities of the sector in which they operate. In fact,

⁴⁷ [Trusted Data and Artificial Intelligence Systems \(AIS\) Playbook for Finance Initiative.](#)

⁴⁸ <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

the use of such list and the surrounding principles and requirement set by Ethics Guidelines for Trustworthy AI, combined with the Ethically Aligned Design principles, are key for the operationalization of ethical mandates in the financial domain towards the responsible and sustainable AI innovation in the financial sector, helping the actors involved to better understand the risks that the AI system might generate, and how to minimize or avoid them, whilst maximizing the benefit of AI applications.

The ALTAI list lingers over seven ethical principles for Trustworthy AIS and is deeply grounded in the protection of people's fundamental rights enshrined in the EU Treaties, the Charter of Fundamental Rights (the Charter) and international human rights law: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being; and accountability.

ALTAI principles can be usefully exploited to map the AI-driven use cases/applications with the ethical implications relevant for the requirements provided by the Ethics Guidelines for Trustworthy AI. This exercise has been already conducted in relation to twenty **AIS Use Cases for Financial Services**, which were observed throughout the global financial services industry and relying on the massive increase in the development, use, and implementation of AIS applications. These kinds of applications have supported and are going to further empower financial organizations in delivering incremental value through several avenues, such as engaging customers, reduced credit loss, increasing share of wallet, customer acquisition, and improved profitability. At the same time, they imply ethical concerns. The use cases were classified in the following four business categories: Product and Customer; Risk; Operations and Corporate: for each of the use cases major ethical concerns need to be considered.

For instance, in the category "Product and Customer", the use case "Personalized Marketing Offers" allows institutions to target customized promotional material to individual customers, thanks to AI models trained on customer marketing behaviour. Such models are used to adapt the marketing materials and thereby better allocating marketing budget in relation to the customer propensity, as well as determining the most appropriate distribution channel. The main ALTAI-driven ethical implications in this case regard Privacy and Data Governance, Diversity, Non-discrimination and Fairness, Transparency, Societal and Environmental Well-being.

In the category "Risk", the use case "High-frequency Trading/Robo-Advisors" is related to AI-based portfolio management services. In particular, it consists, on the one hand, in the high-frequency algorithmic trading. This has completely changed market dynamics: the trading is executed by fully autonomous AI systems (without human intervention), thanks to the ML system which analyses, recommends, and executes future trades on the basis of all data from every trade captured and fed back into such ML system. On the other hand, the use case refers also to the robo-advising service. This service is based on similar trading algorithms applied to less frequency trading executions, at both the institutional and personal portfolio management. In other words, personalized investment recommendations are offered by ML which replace human-generated trading strategies. The service ranges from a basic level to an advance one. In the form, the trades are automatically executed thanks to the preselection of a type of trading strategy backed by an AI-based trading algorithm (given a set of investment guidelines). The latter foresees a greater access to customer risk preference data for training the AI system, to provide hyper-personalized investment recommendations. The main ethical implications pertain to Human Agency and Oversight and Societal and Environmental Well-being, in particular as regards the potential workforce displacement. It is necessary to ensure a people-first approach when designing and adopting the application.

From a wider perspective, it is paramount that in the financial sector the **socio-technical systems “remain human-centric, serving humanity’s values and ethical principles”** and “operate in a way that is explicitly beneficial to people and the environment”, in addition to reaching functional goals and addressing technical problems. This approach will foster the heightened level of trust between people and technology that is needed for its fruitful use in our daily lives”⁴⁹

The **ever-increasing number of ethical resources** developed by academics, consulting firms, technology firms and other organizations and institutions include principles, codes of conduct, guidelines, White Papers and research papers: this wealth of content can be overwhelming and many of these documents can be not relevant for the financial domain.

For developing, executing, and scaling Trusted Data and AIS, it is necessary to focus on **three critical building blocks**:

- **People**: it is key to focus on developing a strong data and AIS climate and culture, in conjunction with a clear data and AIS strategy and education and awareness on the trusted data and AIS initiatives;
- **Process**: it is paramount to have in place “a series of data and AIS ethics governance tools including principles, data, and AIS impact assessments, as well as measurements and KPIs. Standards and certifications are also starting to appear (i.e., IEEE std 7010TM, ECPAIS), which will allow organizations to certify both their technical and governance initiatives for trusted data and AIS”;
- **Technology**: it is necessary to have the appropriate technological solutions to scaling trusted data and AIS projects, leading to the successful implementation of value driving AIS projects. This technical process “is platform agnostic and... includes modern DataOps, differential privacy, a feature farm, AIS and analytics models, trusted MLOps, business decisions and value calculations, testing and optimization, and modern DevOps”.

With respect to the financial sector, active interest and concrete steps for regulating AI globally have been showed mainly by the **two types of regulators**, focused on different subjects: the Privacy Regulators, focusing on consumer privacy and data protection matters, and the Prudential and Securities Regulators, which regulate banking and financial market matters.

As for the former, the GDPR in the EU is recognized as the leading regulatory source in terms of privacy including AI-related issues, serving as a model and influencing other lawmakers to draft similar legislation to protect personal data.

As for the latter, in terms of prudential and securities regulatory oversight, the Monetary Authority of Singapore (MAS) in 2018 published the FEAT principles (Fairness, Ethics, Accountability, and Transparency), which are now renowned worldwide. They set a first high-level framework for financial institutions when instating AIS governance measures.

It is also significant to mention the identification of the **top upcoming regulatory topics** that are expected to be emphasized and regulated in terms of legislative amendments and regulatory requirements, besides those stemming from privacy legislation:

- **Transparency**: organizations should be committed in ensuring transparency while disclosing relevant information to their constituents, including the disclosure to data

49 Introduction of the Ethically Aligned Design, First Edition.

subjects when they are the subject of a decision based exclusively on automation using personal data. Furthermore, automated decision making should be monitored by human employees, and it would be advisable to foresee the right of consumers to opt-out of decisions made solely by machines and request for human decision making. The existence of AI and automation should be disclosed to the constituents, with a right for the data subjects to “be informed when interacting with AI applications”;

- **Explainability**, which, broadly speaking, is the right to request an explanation for who and how the decision is made. There should be obligations regarding explainability for any decisions that are made exclusively automatically, such as the obligation to inform data subjects of personal data that was used in automated decisions and all the factors that led to a decision. Some of these requirements apply only when the decision rendered is exclusively based on automated processing. The concept of “exclusively” should be interpreted as process that is totally automated and excludes any human influence on the outcome.
- **Fairness**, for which reference could be made to Canadian legislation, where fairness is largely governed by anti-discrimination legislation, rooted in protecting equality rights and could be enriched by a Human rights-by-design requirement;
- **Accountability**, including end-to-end accountability and data traceability. It implies organizations to document model development and monitor the models afterward, as well as privacy practices. Its execution encompasses, for instance, “information such as where data is originally from, how it is collected and curated, where it moves in the organization from its source to its destination, how the data gets transformed, where it interacts with other data, how it is prepared and labelled, having audit trails recording the correlations and inferences made algorithmically in the prediction process, how data accuracy is maintained over time”.

Among other interesting narratives exploring from an ethical perspective the risks posed by the wide scale adoption of AI in financial services, the following can be mentioned:

- The report “**Navigating Uncharted Waters: A Roadmap to Responsible Innovation with AI in Financial Services**”, developed by the World Economic Forum and Deloitte. It identifies areas of regulatory uncertainty and outlines insights and findings on the concerns and risks of AI to the financial system, such as the AI explainability, fairness, algorithmic bias, fiduciary and collusion. Among the main risks, it mentioned: i) the risk of customer backlash from AI failures, causing reputational damage; ii) the risk of censure, regulatory scrutiny or depletion of goodwill; iii) the risk of alienating employees; iv) the systemic risk posed by AI which raises new questions on financial stability and new sources of system risk, such as the herding to move markets, human panic due to the unpredictability of machines, and the effects of optimizing algorithms locked in competition (e.g., rate setting); v) the perpetuation and/or exacerbation of unfair biases (human, data, model, second-hand) in financial decision making; vi) the risk of not meeting various standards of fiduciary duty caused by the expansion of AI systems’ responsibility; vii) the risk of erosion of the traditional financial systems’ defences, including the undermining of regulatory control mechanisms and the loss of skills in human, who are removed from the process. The document also seeks to propose mitigating measures to address key concerns surrounding the use of AI in financial services in order to contribute to safely unlock the power of AI in such sector. These measures range from exploring the responsible and trust-first AI business models of the future, to the responsible deployment of AI systems through new models of governance to challenge the foreignness of AI;

- The document elaborated by the Monetary Authority of Singapore, relying on the consultation with financial services organizations: **“Principles to Promote FEAT in the Use of AI and Data Analytics in Singapore’s Financial Sector”**. The document highlights 14 high-level principles, aimed at strengthening the ethical use of AI and data analytics (AIDA) in financial products and services, providing foundational guidance on ethical values and governance. The adherence to such principles is expected to contribute to the development of public trust in the use of analytics and AI in Singapore’s financial services sector. The principles are grouped under four sections where, for each principle, a brief explanation is followed by several illustrative examples specific to financial services. The sections are:
 - Fairness, covering the justifiability of the decision-making process. It encompasses the use of personal attributes, as well as accuracy and bias, and calls for regularly scheduled reviews and validation.
 - Transparency, which envisages several measures to ensure public trust, such as the disclosure of the use of analytics and AI to data subjects, as well as the explanation (upon request) of the data used and the effects of decisions made by analytics and AI systems.
 - Ethics, which underlines that the use of AI and analytics is aligned to other organizational code of conducts and ethics standards, besides calling for AI and analytics to be governed by the same ethical standards as human decisions.
 - Accountability, including both an internal dimension and an external one. As for the former, it is recommended that all uses of AI and analytics are approved by an internal authority and that the management and the board of directors are informed of all use cases. As for the latter, data subjects should be empowered with a feedback loop to enquire about and/or appeal any decision made by AI, besides the obligation for any company to verify any relevant data provided by the subjects in their decision-making process.
- A high-value soft-law source from Canada, the **“Code of Conduct for the Ethical Use of AI in Canadian Financial Services”**. This document contains a set of principles, again developed in consultation with several Canadian financial services organizations. It is directed to foster the ethical use of AI in financial institutions/organisations, by providing practical guidance to prevent ethical implications in their day-to-day use of AI. It is very interesting source, since it represents the first step towards a granular, practical, sector-specific set of guiding ethical principles. Besides a series of definitions, the document lists 17 principles, together with practical examples, grouping them into the following three categories:
 - Principles of Fairness: any kind of disadvantage of individuals generated by bias and discrimination, both in an intentional or in an unintentional manner, needs to be avoided through ethics reviews for material AI applications. This group of principles also covers justifiability and the need to extend the lawfulness paradigm with the practice to understand the aggregate input data and its outcome before using it for material AI applications.
 - Principles of Accountability, which calls for the need to identify specific individuals in the financial organization/institution as responsible and accountable for the output of such organization/institution’s AI applications, regardless of where they are developed. They should be responsible for the level of autonomy assigned to an AI system and should be held to the same standard as human decisions. Besides this aspect, the document outlines

the need for approval of any AI application by the accountable and responsible parties prior to production, in order to certify that a consistent model development process has been followed and to ensure that such model is validated and undergoes a monitoring process;

- Principles of Transparency, which encompasses explainability, especially in case of models making decisions that affect individuals. Any data used in the AI application should be compliant with the highest standards of data privacy and informed consent and the financial institution/organization should timely inform individuals of their direct interaction with the AI application (i.e., chat bot), or when a decision is taken by an AI application (i.e., credit lending).
- The report elaborated by the Berkman Klein Center for Internet and Society at Harvard University: “**Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI**”. It is the outcome of a survey on 36 prominent ethical AI documents, conducted with the aim of assess, identify and summarize key themes covered by all of them. Though not specifically focused on the finance sector, its findings are relevant for it, also considering that it highlights that the ethical principles are context-specific and need to be interpreted in their “cultural, linguistic, geographic, and organizational context.” The key principles identified are privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility, and promotion of human values.

It is also important to mention the **emerging standardization initiative and certification programmes supporting the Ethically Aligned Design**: they can allow financial organizations to certify their commitment for trusted data and AIS.

In particular, one of the leading projects in this regard is the **IEEE P7000™**, which establishes a set of processes for ethical values elicitation and prioritization throughout the end-to end-lifecycle.

Although it was not specifically developed for the financial sector, it can be valuable for ensuring ethically-inspired development of AI, data and analytics solutions in such domain, providing traceability of ethical values to embed value dispositions in the system design.

IEEE Std 7010™-2020, on the other hand, might be relevant to the financial sector, referring to the well-being: it involves a comprehensive view on what brings a person physical and mental health, as well as the environment, thus it might help to move determining a holistic value in the algorithmic age.

As for certification, it is recommended to explore the applicability of the IEEE-approved **Ethics Certification Program for Autonomous and Intelligent System (ECPAIS)** to AI-based applications, systems, services and solutions in in financial domain, functional to promote responsible innovation, which, in turn, is expected to enhance the confidence in the financial organization/institution concerned: the design and implementation of a solution against ECPAIS (Ethics Certification Program for Autonomous and Intelligent Systems) criteria demonstrates the commitment towards a trustworthy foundation of the AI system.

3.2 DEPLOYMENT OF APPLICATIONS AND REAL-LIFE SCENARIO RECOMMENDATIONS

The deployment and experimentation of applications is pillar in the innovation. The challenges faced in this topic are many, and the suggested approach is to test in a safe environment, to use cloud-based testbeds and sandboxes, and to rely on the DevOps approach to integrating the development task (Dev) with the operation tasks (Ops) for the deployment and maintenance of the applications in the private or public cloud. DevOps is the emerging approach to agile software development that developers and operations teams use to build, test, deploy and monitor applications with speed, quality, and control. Compared to other industries, DevOps is starting to become essential for the FinTech due to the unusually high priority given to core aspects like maintainability (FinTech applications require improved analyses with massive amount of data, new features and data integrations must roll out very frequently and with very low latency), quality (an uncaught mistake or security hole quickly runs to huge monetary losses or fines, not to mention the reputability of the institution), data integration (FinTech applications process a continuous and never-ending stream of new information, and the more feeds the better), transparency and traceability (FinTech applications need to support forensics activities to spot fraud, extortion, money laundering, and other criminal activities).

The development of a scalable and flexible architecture allows an eased deployment of applications, especially relying on adaptable building blocks.

An example of a concrete realization of many of the previous concepts and approaches has been produced within the context of the INFINITECH project ⁵⁰.

Starting from the Logical view of the INFINITECH Reference Architecture (IRA), with a full alignment to its building blocks, the IRA Development and Physical views have been tackled and designed in terms of the concrete specification and realization of the fundamental and target INFINITECH concepts of Testbeds, Sandboxes and Datasets management, and related tools and techniques for their effective setup and deployment in the INFINITECH FinTech and InsurTech pilots and validation real-life scenarios, focused on topics like:

- Smart and Reliable Scoring, Risk and Service Assessment;
- Personalized Retail and Investment Banking Services;
- Financial Crime and Fraud Detection;
- Personalized Usage Based Insurance Products;
- Configurable and Personalized Insurance Products.

In the INFINITECH context, a pilot *TESTBED* is defined as the set of resources like compute, storage and network that support one or more pilot use case applications. Figure 12 below represents this graphically.

More specifically, a testbed can be considered from a physical view (the actual configuration of the aforementioned physical resources). The resources that compose a testbed can be inside a private Data Centre or in any cloud provider.

⁵⁰ <https://www.infinitech-h2020.eu/the-project>

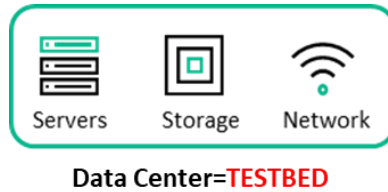


Figure 11 - Data Center (INFINITECH Project)

Moreover, a testbed can be considered from a logical view (as the set of virtual resources managed by an orchestrator like Kubernetes (K8s), an open-source system for automating deployment, scaling, and management of containerized applications). In particular the capabilities of a *K8s cluster* are leveraged to manage the testbed. In the specific case of an INFINITECH testbed hosted on a cloud provider like AWS (Amazon Web Services), the K8s capabilities are provided by the EKS (Amazon - Elastic Kubernetes Service) service.

With this approach in INFINITECH it is possible to manage a scalable number of pilots and testbeds like depicted in the following figure:

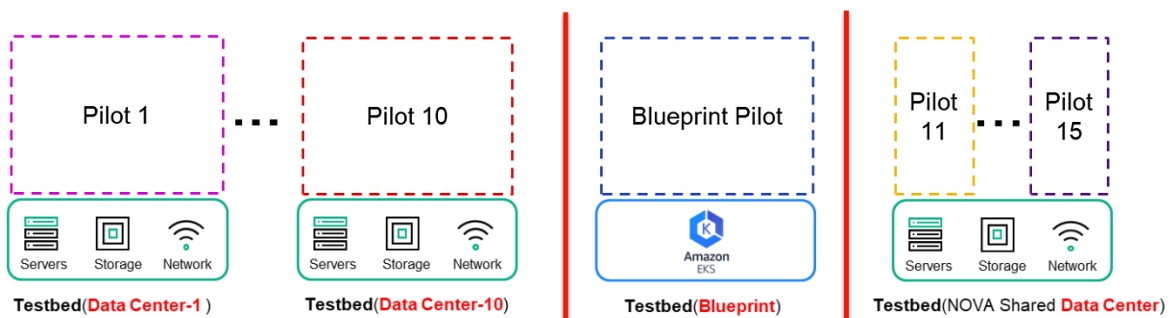


Figure 12 - Testbeds (INFINITECH project)

Within INFINITECH pilots' deployments, the concept of *SANDBOX* is also relevant. From a general point of view a sandbox can be defined as an isolated testing environment that enables users to run programs or execute files without affecting the application, system or platform on which they run. In the INFINITECH context each pilot has one or more use cases (realized by one or more pilot applications, each one typically realized by one or more INFINITECH microservices): in this case each use case is a Sandbox provisioned by the leverage of *K8s Namespaces*.

In fact the K8s Namespace feature makes it possible to logically isolate the objects (mainly POD – Point of Delivery) inside it from other Namespaces. Therefore, each dedicated testbed will only have one K8s cluster with as many Namespaces as the number of use cases to be implemented for a single pilot. In the case of a shared testbed, there will be one K8s cluster for each pilot it has to host and manage. In the end, in this general context, each pilot App will be realized as a K8s POD. The following pictures depict the previous concepts.

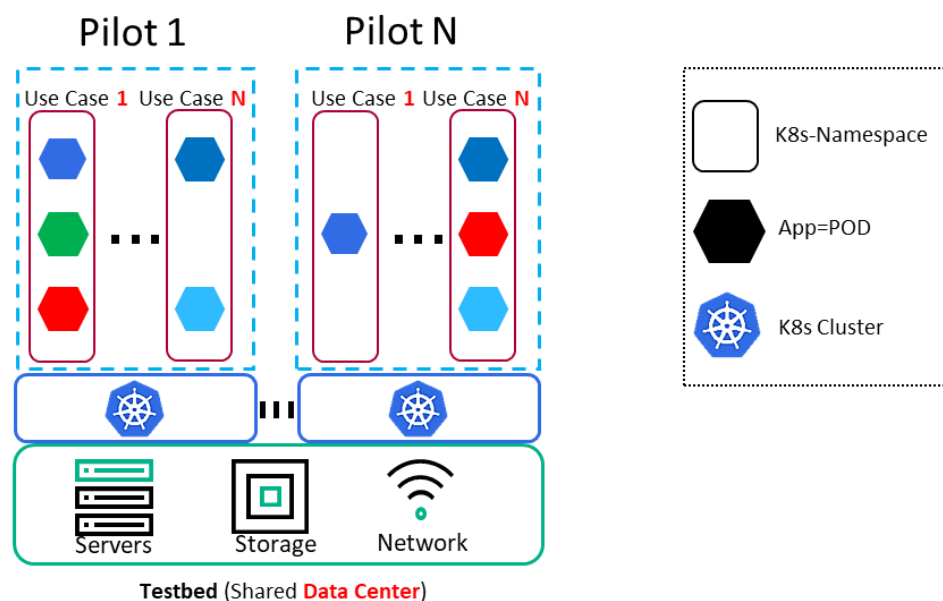
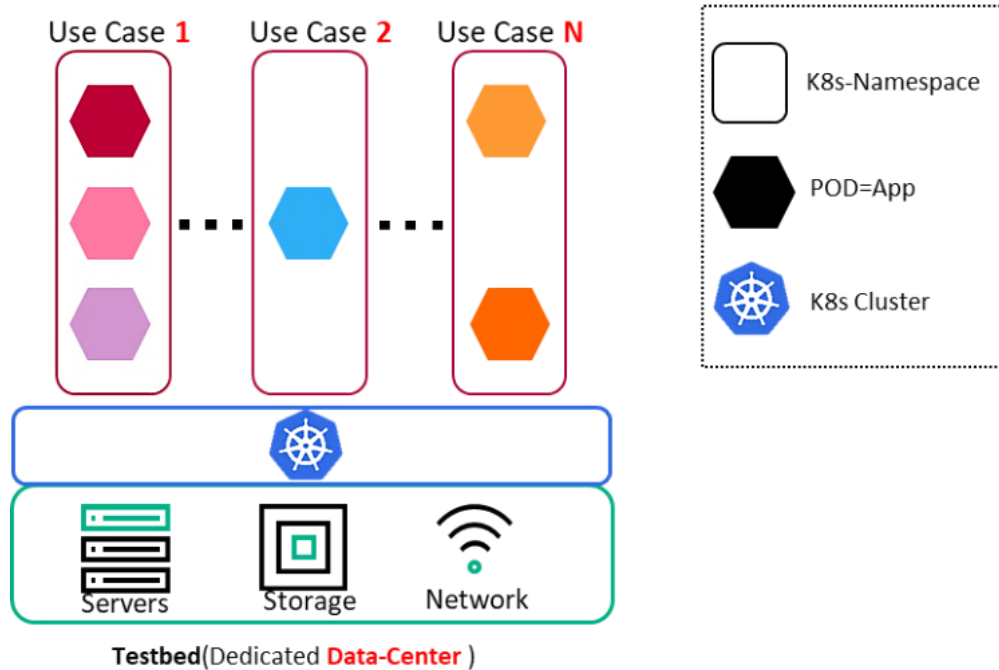


Figure 13 - Data Center, Testbed and Use Cases (INFINITECH Project)

Finally, in order to fulfil the INFINITECH project goals and effectively manage pilots' testbeds and sandboxes, it has been decided to implement DevOps processes, whose practical implementation goes through the Continuous Integration / Continuous Deliver (CI/CD) processes. The CI/CD processes have been created in the context of a blueprint reference testbed environment: with this approach, it is then easy to spin up new testbed environments from scratch, which enables future scenarios including automated end-to-end integration testing.

Moreover, it is already foreseen to enhance the processes by adopting a DevSecOps (Development, Security and Operations) approach and including the related tools in the CI/CD pipelines, which will enable to include security in the testbeds and sandboxes software development and deployment life cycle from the beginning, following the same principles of DevOps. Security is then considered throughout the process and not just as an afterthought at the end of it, so that different kinds of security checks are executed continuously and automatically.

3.3 THE EUROPEAN UNION MARKET: THE MAJOR AI-FINTECH HUB OF THE WORLD

One of the critical factors for the consolidation of a true ecosystem around AI to improve the design and delivery of financial services in the EU is undoubtedly the attraction and retention of talent. While it is true that in recent years various initiatives have been developed at the national level and a project driven by the EC at the European level⁵¹ the European ecosystem has not yet consolidated its position in the global market.

The bet, to date, seems to be the same as always when it comes to the EU: each country creates its own infrastructure, arranges its priorities on the basis of a national policy and injects funds in a haphazard way to enhance the growth of its local nodes. While it is true that most of the initiatives have been promoted by the private sector: Portugal Fintech, Innsomnia, Fintech District, Holland Fintech, Copenhagen Fintech, Frankfurt Accelerator⁵², among others, there are also initiatives that are directly rooted in the public sector and from which services are provided to both the regulator and the national authorities: there are the cases of The Lhoft or Fintech Scotland.

Since 2017, The Talent Route⁵³ brings together all these national initiatives to give them a European dimension and thus try to compete with the most powerful AI-Fintech ecosystems in the world: Asia and the United States. In this sense, it is now more necessary than ever to join forces and encourage collaboration between all the players in the European Fintech ecosystem. It is worrying, as Luis Moreno and Andrés Pedreño point out in their book “Europe versus the United States and China: Preventing decline in the age of artificial intelligence”, that the EU has “ceased to be a benchmark in innovation to become the passive object of the greatest global transformation process ever seen by human beings”, and, in that sense, it seems urgent to orchestrate a strategy that focuses on the following strategic axes⁵⁴.

- 1) It is necessary **to unite all the accelerators and Fintech hubs in the EU** to maintain talent and attract new capabilities from other global ecosystems. With that purpose, projects such as INFINITECH, can serve as a “global test” by allowing the configuration of a virtual space of connection between all agents.

51 https://ec.europa.eu/info/publications/180620-eu-fintech-lab-meeting_en

52 At the moment , 13 Hubs from 10 EU countries are part of The Talent Route. The websites of all these organizations can be found at: www.thetalentroute.com

53 www.thetalentroute.com

54 These goals are in line with the EU's Fintech Action Plan, which has established three priorities, namely: supporting innovative business models and their scaling, reinforcing the adoption of new technologies in the sector, and promoting a system of integrity, ethics and security in relation to financial services.

More information about this plan can be found at: https://ec.europa.eu/info/publications/180308-action-plan-fintech_en

INFINITECH VDIH (Virtual Digitalized Innovation Hub) aspires to become, in collaboration with The Talent Route network, the largest specialized Fintech hub in the world. To achieve this, the following steps envisaged in its action plan must be consolidated:

- Identification of all HUBs/Accelerators operating in the EU, both public and private.
 - Mapping of the main EU Fintech solutions, particularly in the AI framework.
 - Design of a catalogue of services offered by each hub to design a joint service catalogue based on specialization.
 - Design of a digital work platform that allows the development of PoCs and the connection of test environments.
 - Launch of an IA-Fintech Call.
 - Creation of the Fintech Campus in a pilot network of universities.
 - Connection of the VDIH with other hubs worldwide.
 - Development of the 100 Startups event at European level.
- 2) The EU must create a **map of specialization by Fintech technologies and competencies around AI** that allows us to identify the potential growth of our offer and its investment needs, training and connection with large operators (mainly Corporations). In that sense, Insomnia publishes annually its Galaxy Fintech ⁵⁵ where the 1000 most important Fintech initiatives at European level are collected.
- 3) **The EU is the largest Fintech laboratory in the world:** we develop more solutions than any other ecosystem, probably because our market is very fragmented and, around each hub, entrepreneurship opportunities are developed, but our projects are small (1 to 10 workers and 500K € turnover on average) and this is not scalable, therefore, it is more necessary than ever **to create more favourable investment environments and more flexible rules that favour the aggregation of local initiatives.**
- 4) **AI as a backbone technology:** according to the latest Summit of The Talent Route, more than 40% of the Fintech projects being developed in the EU are based on AI. Ambitious projects being developed from AI in fields such as scoring, the fight against fraud, forecasting delinquency or predicting the behaviour of companies. It is clear that we have our own talent capable of reflecting on many ideas at the same time. For this knowledge to consolidate and grow, it is crucial to establish transfer mechanisms (lack of symposiums, comparative studies, joint projects) and instruments for connection with the academic world: universities must urgently implement specialized training in AI that subsequently connect with industry.

55 <https://insomnia.es/es/fintech-land.html>

- 5) **Test bed⁵⁶**: this network of Hubs and Accelerators must weave an ecosystem in which, in addition to start-ups and scaleups, financial institutions, national banks and the regulator, both national and European, must play an important role. Once it has been proven that we are leaders in early-stage entrepreneurship, what we need to scale, in addition to capital and size, is to be able to test fast and fail cheaply. The test environments (sandbox) are essential for initiatives to make the leap or to market earlier and, thus, can convince investors faster.
- 6) **Ethics and legality: we cannot fail to consider ethics as a consubstantial part of all AI solutions that are developed in the EU**, therefore, it is essential that there is, at source (in schools), a basic training on the concept and scope of AI: the EU would like to spread to the largest audience possible within the Member State a basic knowledge of AI that allows the development of a market that generates confidence, along with clear boundaries imposed by regulations and legal environment particularly with regard to the GDPR.
- 7) **Global vision: AI-Fintech solutions from the EU have to aspire to compete in other global markets**. Once tested, the solutions must be internationalized. For this, it is important that the EU designs *ad hoc* internationalization policies and uses its networks and bilateral agreements for this purpose in the coming years. In addition to the Asian and US markets, LATAM and the African region represent a great market opportunity in the short term that we are not taking advantage of because we are not able to design an ambitious strategy around AI in both markets, namely LATAM and African.
- 8) **Digital Finance Academy**: as a result of all these movements we will be able to provide the European Academia with empirical contents. While it is necessary that we urgently need to incorporate all these reflections in Universities (hackathon models and the development of public-private collaboration initiatives can serve as methodologies to connect students for their training in new skills), it is no less important that the Digital Finance Academy must also be based on the vision of market problems and the development of solutions that come from practice.
- 9) **The AI-Fintech call**: in order to achieve the largest scouting exercise in the history of the EU, a large call should be designed in which the main financial institutions in the EU define the challenges of the future of AI in relation to financial services. From the reading of this exercise, moreover, the outlines of the industry's challenges for the next two years could be drawn.

⁵⁶ These goals are in line with the EU's Fintech Action Plan, which has established three priorities, namely: supporting innovative business models and their scaling, reinforcing the adoption of new technologies in the sector, and promoting a system of integrity, ethics and security in relation to financial services.

More information about this plan can be found at: https://ec.europa.eu/info/publications/180308-action-plan-fintech_en

4 INNOVATIVE TRENDS AND CHALLENGES IN APPLYING BIG DATA AND AI IN DIGITAL FINANCE

Digital Finance brings innovation to the Financial Sector by either improving existing solutions or disrupting role and structure to develop cutting-edge applications. The several technologies utilized in the financial environment follow different trends and the accelerating pace of technological change is the most creative force in the ecosystem today. Specific relevance is given to AI and Robotics and Big Data, however other innovative trends will be mentioned and briefly described to allow a more comprehensive overview of the topic.



Figure 14 - Financial Technology Space, Diagram by T. Butler

The objective of this White Paper is not to deeply describe the technologies present in Digital Finance solutions, rather to provide an overview of the trends and challenges of their applications, focusing on Big Data and AI. Indeed, this section is intended to bridge the gap between research and industry by showing the common interests and highlighting the synergies.

Annex II contains a collection of use cases representing interesting applications in the Digital Finance Domain, the use cases are split into categories and summarized below:

Banking Sector: the following use cases address pains in the Banking Sector exploiting AI-based solutions. Some interesting business cases along with their expected outcomes can be found thoroughly detailed in the Annex.

- **Intelligent Document Processing (IDP):** The intelligent document processing solution is leveraging Natural Language Processing and Deep Learning in order to classify documents and extract relevant information automatically.

- **Data Driven Decisions (3D):** The 3D solution aims to analyse risk management frameworks along with undetected patterns to determine the Probability of Default of banks' customers loans.

Financial Services: such services comprehend as clients the different players present in the financial sector. Hereby, two use cases resolve different business cases by leveraging mainly on blockchain technologies and a third one proposes a completely new solution to support SMEs.

- **Blockchain in the Digital Asset Supply Chain.** The Blockchain-based solution is designed to bring a broad range of benefits to complex supply chains, including traceability and provenance of physical and digital assets, improved end-to-end visibility and applications of AI and advanced analytics.
- **DAOs in Decentralized Finance (DeFi).** DeFi provides permissionless access to a variety of financial services, eliminating intermediaries and enabling transparency while ensuring privacy.
- **Network Finance.** The objective is to create on a proprietary platform a network of companies to demobilize their commercial receivables without recourse to credit. A new prospective is to reorganize the single one to one relations between debtor and creditor as a network view one to many.

Insurance Sector: the footprint of AI in the gathered use cases is strong, as the first mean to resolve real problems in real scenarios. Both use cases are related to insurances for vehicles, whilst following different approaches to develop different solutions.

- **Personalized insurance products on IoT vehicles.** The solution, based on real data collection using IoT, AI, and cloud computing, develops two business services: (i) Pay-as-you-drive model based on drivers' profiles, and (ii) Fraud detection, that exploits real-time car and traffic data to depict real scenarios.
- **Fraud detection and prevention in car insurances.** This use case relies on white box AI-based algorithms to detect and explain frauds, aimed to automate the detection of frauds relying on antifraud rules and new implemented rules.

RegTech: the following use cases leverages on technologies such as AI to monitor flows, assess processes, and provide useful tools or protocols to address specific cases explicitly related to regulations, compliance, and documentation.

- **Automated Regulatory Risk Intelligence.** The focus in this use case is to extract signals from news/online discussions using natural language processing to identify features of interest, and ML to classify categories of interest and monitor for early identification of topical trends.
- **Harmonised Compliance Protocol.** The objective of the use case is to develop a protocol to classify financial institutions' clients in accordance with applicable regulation and determine whether a financial institution can on-board a client to trade with it in a particular service, product and country whilst performing appropriateness and suitability assessments.

Others: the last use case, as well as the last categorization, shares some components with other use cases listed above, still addressing a different and specific segment of the market: trade investor information.

- **Audio Experience from Investors Reports.** This solution is based on an AI/ML-based text-to-speech solution, which transformed documents in formats which lack sufficient metadata and rich mark-up languages into audio experiences to be accessed through a mobile application and a complimentary web app.

4.1 AI AND ROBOTICS IN DIGITAL FINANCE

The term AI comprises different concepts, that could be classified into three logical levels.

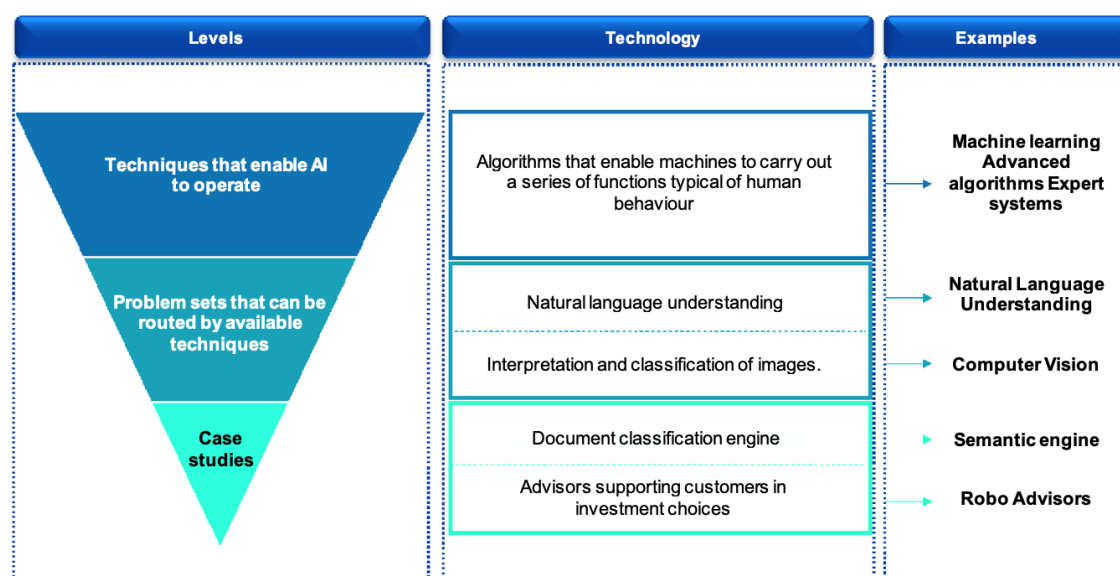


Figure 15 - Guide to understanding artificial intelligence , source: ABI Lab

The first level features a series of AI disciplines, such as ML, advanced algorithms, and expert systems, which create specific algorithms to enable the systems to carry out a set of functions such as learning, understanding and interpreting a context, or making automatic decisions.

The intermediate level includes the classes of problems that can be run through AI, e.g., **Natural Language Understanding (NLU)**, or **Computer Vision**.

The case study level features real business cases that can be run and handled through AI (e.g. **Chatbots**, or **Robo-advisors**).

The banking sector is at the frontline in assessing and understanding how AI can fit into the ecosystem and favour digital transformation processes.

The current scenario shows that, following a period of scouting or isolated experiences and testing, the projects based on AI technologies are being included in banks' business plans on a more structured basis, along other initiatives already underway or planned.

As part of the survey on ICT priorities, ABI Lab investigated the application of AI in the banking sector to identify the existing and emerging trends.

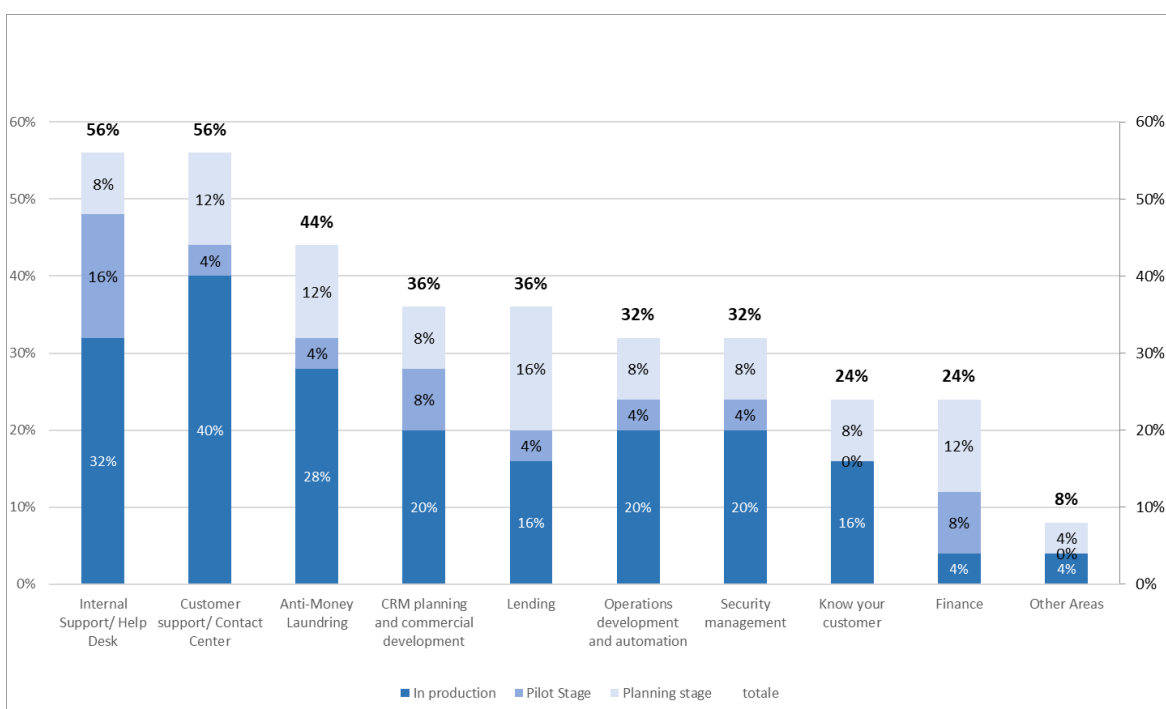


Figure 16 - Application of AI, source: ABI Lab survey on ICT priorities

The chart above shows that AI is mainly used in the areas of **internal and customer support, commercial planning and development and operations**.

The **Anti-Money Laundering area** represents a quite unique case, in fact despite the small percentage of application (20%), all cases examined were already in production.

Lastly, Lending, Finance and Know Your Customer (KYC) are the areas with the smallest cases of AI in production.

The other areas reported in the chart include the support to compliance to detect the accuracy and entirety of documents.

Between the specific use cases it is also possible to mention the development of **Automated consultants**, often known as robot advisors, that are considered among the most interesting application of AI in financial services. The European Supervisory Authorities joint report (2017a) defines the phenomenon of automation in financial advice as “a procedure in which advice is provided to consumers without, or with very little human intervention and with providers relying on computer-based algorithm and/or decision trees.” In practice, robot advisors build personalised portfolios for investors, on the basis of algorithms that take into account investors’ information such as age, risk tolerance and aversion, net income, family status. Obtaining this information is a legal requirement and robot advisors employ online questionnaires to obtain it.

4.1.1 Ongoing challenges

In order for the hype surrounding AI to consolidate into a real catalyst for the transformation of banking, it seems increasingly important to focus on the main challenges with respect to internal operational factors (e.g., project management, the impact on risk mapping and cultural transformation) as well as to ethical aspects and transparency.

Project Management

To favour the adoption of AI in banks, it is crucial to promote the research and testing of the various techniques available, as well as to assess the effects of the technology on the operational and business aspects of the bank.

A **gradual approach** is therefore recommended.

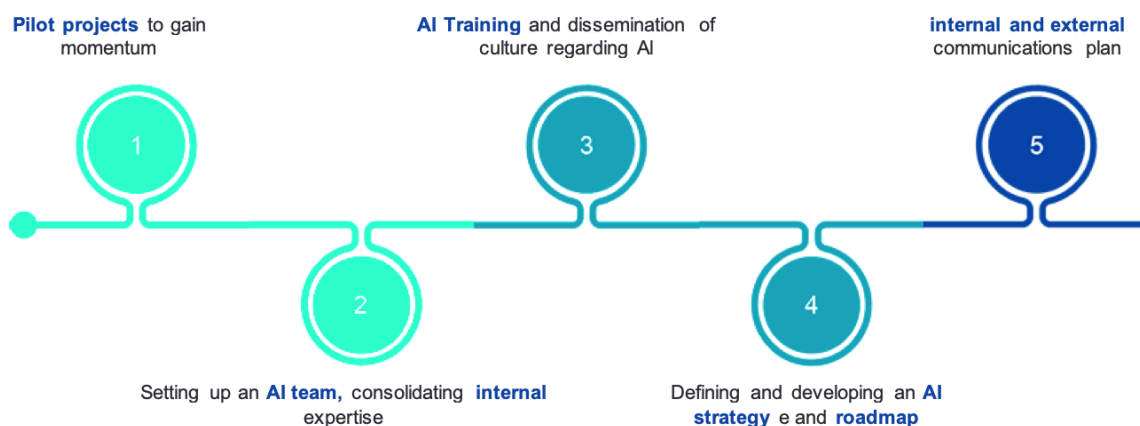


Figure 17 - Phase of an AI project, Andrew Yan-tak Ng

As a result, management must focus on strategies of listening and testing. It is also crucial to choose the right methodologies to adopt in designing the innovative processes enabling the use of AI systems. Such technologies applied to project management practices would improve automation of standardized processes enabling resources to focus on more niche matters enabling as well greater engagement.

Impact on risk mapping

AI can have a significant impact on specific company processes. As a result, it may be necessary to introduce new methods of managing controlling activities. In light of this, it may be useful to:

- assess the impact that AI systems could have on defining the risk mapping (specifically on operating risks);
- set up surveillance areas and monitor possible new risks deriving from the automation of certain activities using AI;
- think of AI as a driver to mitigate certain operational risks, reducing the probability of human error or enabling greater control thereof.

The above list is not exclusive, and it is just a preliminary foundation to drill-down on the topic in the next versions of the paper.

Cultural transformation and impact

As occurred in the past, the introduction of a new technology such as AI brings significant change in people and professions.

However, to correctly assess the impacts of AI, it will be important to focus on several considerations on the technology and the consequences of its use:

- **creation of new professional opportunities and economic advantages:** it is reasonable to assume that AI will contribute to increase jobs, while favouring the creation of new areas of employment, new fields and opportunities. According to the World Economic Forum Future of Jobs Report⁵⁷, 85 million jobs will be replaced by machines with AI, but 97 million new jobs will be created by 2025 due to AI;
- **increasing in worker productivity:** AI may be capable of expanding the tools available to people and their results, due to the increased analysis and processing capacity (augmented intelligence);
- **new ways of working:** it is important to emphasise that man cannot be fully replaced by the machine: the technology will act as a support for people, who, as a result, will have to change their way to work;
- **development of new skills and competencies:** there is a need for continuous development of (technical and management) skills and openness to change, resulting in new competencies to control, enhance and exploit AI-related activities.

4.1.2 EXplainable Artificial Intelligence (XAI)⁵⁸

As AI becomes more pervasive in both our personal and professional lives, more and more stakeholders are interested in understanding AI decision-making. Although AI has grown in sophistication over the last thirty years, the inner workings of intelligent machines are as yet far from transparent.

We may agree that AI decision-making is as opaque as human decision-making. Humans continuously engage their senses to recognize images, sounds, and languages. Yet, humans find it especially hard to provide explanations for the tasks they performed. For instance, when a person is asked to explain the reason, he/she is seeing a cat in a picture, the same person is typically not aware of the specific reasoning involved in his/her decision-making.

When an explanation is requested, humans tend to use a logical method (e.g., pointing out cat's whiskers and ears) which is different from the intuitive method we use to recognize the cat in the very first place. Similarly to humans, ML models are referred to as *black-boxes*, in which the decision-making processes are often opaque.

57 <https://allwork.space/2021/11/ai-will-create-97-million-jobs-but-workers-dont-have-the-skills-required-yet>

58 Paolo Giudici, Emanuela Raffinetti, "Shapley-Lorenz eXplainable Artificial Intelligence", Expert Systems with Applications, Volume 167 (2021) <https://doi.org/10.1016/j.eswa.2020.114104>. Bussmann, N., Giudici, P., Marinelli, D. et al. Explainable Machine Learning in Credit Risk Management. Comput Econ 57, 203–216 (2021). <https://doi.org/10.1007/s10614-020-10042-0>.

In addition,

- i) as regards the impact Covid 19 on digital finance: Agosto, Giudici. (2020) A poisson autoregressive model to understand COVID-19 contagion Risks, 8(3), 1-8, 77;
- ii) as regards the network models: (2020) Giudici, P., Hadji-Misheva, B., Spelta, A. Network based credit risk models. Quality Engineering 32 (2), 199-211; (2019) Avdjiev, S., Giudici, P. and Spelta, A. Measuring contagion risk in international banking. Journal of Financial Stability, 42, 36-51; Ahelegbey, D., Giudici, P. and Hadji-Misheva, B. Latent factor models for credit scoring in P2P systems. Physica A: statistical mechanics and its applications, 522, 112-121; (2017) Paola Cerchiello, Paolo Giudici, Giancarlo Nicola: Twitter data models for bank risk contagion. Neurocomputing, 264, 50-56;
- iii) as regards digital assets and cyber risk: (2019) Giudici, Abu-Hashish. What determines bitcoin exchange prices? A network VAR approach. Finance Research Letters vol. 28, 309-318; (2020) Facchinetti, S., Giudici, P. and Osmetti, S. Cyber risk measurement with ordinal data. Statistical methods and applications 29 (1), 173-185

The opaqueness of intelligent machines may have negative consequences on firms that are considering integrating AI agents into their operating models. More specifically, here are some of the major complexities that come along a non-transparent AI decision-making:

- Firms will not be able to explain the algorithm's outcomes to their clients, auditors, and compliance teams, if necessary.
- Promising models may not be put into production as firms do not entirely trust AI decisions; therefore, having explainable algorithms becomes keen to enable trustworthy AI.
- Firms may expose themselves to significant risk if they delegate responsibilities to AI systems that are not fully in line with their purposes and internal policies; indeed developing fully explainable and aligned algorithms is central.

The risk is to create and use decisions that are not justifiable, legitimate, or that simply do not allow obtaining detailed explanations of their behaviours. Firms may incur serious financial and reputational losses, if an AI-enabled model unjustly discriminates against a specific group of clients, for instance. AI models should not only use training datasets that are statistically representative of the whole population, but also be able to quickly adapt to changes in business objectives. Therefore, explanations supporting the output of a model are crucial in the financial sector, where employees or stakeholders require far more information from the model than a binary outcome for supporting their operations.

Definition and characteristics

Before discerning the characteristics of XAI it is convenient to establish a common ground on what the term *explainability* entails in this White Paper. *Explainability* refers to the ability of a model to make humans understand its function and its decisions. Such a definition includes not only the degree to which humans can understand how the model works but also the ability to be transparent on its own. In line with literature, we define an XAI as one AI that provides reasons or details to make its functioning clear or easy to understand to a given audience.

Therefore, we can picture an XAI as an explanatory agent that reveals its underlying causes to its or another's decision-making (Miller, 2018.) Ultimately, XAI is a human-agent interaction problem. It is necessary to have techniques to trace causality, select the causes, and eventually provide a shared decision through a rule system.

According to Miller (2018), explanations should have the following characteristics:

- Explanations are *contrastive*. People are not interested in understanding why a specific event occurred, but rather why that event happened instead of another expected event.
- Explanations are *selected*. Humans tend to select a restricted number of explanations from a myriad of different causes. However, this selection may reflect certain biases.
- Probabilistic explanations are *ineffective*. The most likely explanation is not necessarily the optimal explanation for a person. People tend to give more importance to the underlying causes of an event than to statistical inferences.

Explanations are *social*. They are a transfer of knowledge from the explainer to the explained, regardless of the type of language used (e.g., natural language or computational language.)

Objectives and benefits

Financial institutions (FIs) are traditionally reluctant to digital transformations that can expose their assets. It can be argued that over the last thirty years few AI techniques were implemented in the financial industry. This was attributable to the opaqueness of “black-box” models as well as a lack of trust in AI. Explainability can help fill this gap that prevented many FIs from adopting AI techniques.

XAI can effectively bring transparency to AI while enabling humans to understand its decision-making process. As a result, XAI can address pressures, such as regulation and compliance, and embrace good practices around ethics and accountability. There are numerous benefits that firms can derive from XAI and for which investing in explainability is business-wise attractive. As XAI guarantees transparency over vulnerabilities and flaws, AI stakeholders can be assured that the model is operating as desired. Even more, the identification of vulnerabilities and flaws can be used to optimize the overall model. This implies that firms can not only improve their performance but also gain better insights into business drivers.

Design

XAI’s ability to open black-boxes and to untangle the complexity of decision making requires additional software components and design paradigms. XAI is no different from any engineering process in which a system’s functions are set in the design phase. Explainability has to be designed *a priori* to effectively affect the choice of ML algorithms. This implies a deep understanding of the mathematical model behind any prediction executed by the AI agent.

A model can either be **explicitly represented by formulas and rules** or **adopting an inductive reasoning approach** to learn automatically from examples and interactions with a reference environment. The former is also known as symbolic AI, whereas the latter refers to the subsymbolic AI.

On the one hand, the symbolic paradigm has its roots in the study of cognitive sciences which makes it very inclined towards logical deductive processes and object-oriented programming. The symbolic processor transforms symbolic inputs into symbolic outputs using hierarchically ordered logical rules. Expert systems are fundamentally based on the symbolic paradigm which allows them to reach a high level of explainability. Programmers can enter rules in the expert system, which in turn can explain its decision. However, it is clear how tautologic such an approach is. The explanation is first introduced by a human, so it should not surprise that the system provides the same explanation as it was instructed to. The system is simply creating a concatenation of rules or facts and can show the logical reasoning behind its conclusion.

On the other hand, the subsymbolic paradigm has its origin in the study of connectionism in which a model simulates the fundamental neural processes in the brain. Inputs are processed in parallel and mathematically computed by neurons. The subsymbolic AI is most often used for tasks such as classification, image recognition, and language recognition. Unlike symbolic models, neural models are not easily explainable. Subsymbolic AI’s low explainability is due to our difficulty in defining a general topology for different applications and generating training datasets that contain no biases. It can be argued that trainers are fully accountable for the decisions a model makes. For instance, if a bank’s lending process discriminates against different groups of people, a model will present the same biases in its decision. Therefore, trainers must be able to inspect a model’s generalization ability and to investigate the criticality of certain decisions.

Although the symbolic and subsymbolic paradigms differ in characteristics and approaches, the combination of the two can help ensure the interpretability of a decision process. The combination of symbolic techniques (e.g., model checking) with subsymbolic techniques is increasingly necessary to ensure the behaviour of neural networks under certain constraints, which can be imposed by a person who wants to ascertain the decisions made by the machine. In other words, we must still rely on both subsymbolic tricks and logical representation models to open black boxes. In the long run, however, models are expected to be explainable *a posteriori* through model rewriting and impose representations and causal constraints to learning algorithms by design.

Techniques

There are numerous different methods available that can provide explanations. Such selection is often grouped into models that are interpretable by design and those that can be explained by external XAI techniques. The description of each of these techniques is beyond our scope, so a brief introduction of the most popular approaches will follow. The first group of models is also referred to as transparent models because they are understandable by themselves. Decision trees fall into this broader umbrella of models that can easily fulfil any constraint for explainability and transparency.

They are hierarchical structures for decision-making often used by AI experts to support classification and regression tasks. Banks have started to use decision trees to understand and inspect the AI decision-making process behind a credit analysis of a given client. Another example of transparent models is the so-called fuzzy logic which is a mathematical tool based on non-standard sets that allow the definition of rules formulated verbally on inaccurate domains. This technique is generally used to link the symbolic approach to the subsymbolic one, thus guaranteeing the explanation of decisions made by an AI agent. Fuzzy rule systems are capable of ensuring explainability since they operate in linguistic terms. They follow a non-binary but multi-value logic (i.e., the sentences can be more or less true), implying a notion of graduality. Fuzzy models can help banks interpreting under which conditions an individual income is classified as low, medium, or high.

If a model cannot guarantee transparency by itself, then external AI techniques may be adopted to explain a model's decision. Such external techniques – also referred to as post-hoc explainability techniques – aim at transferring useful information about how a model generates some predictions for any given input. The many authors view and differentiate post-hoc explainability techniques that are model agnostic from model specific. Model agnostic techniques are fundamentally black boxes explainers that can be applied to any ML model, regardless of its internal representations. Model-specific techniques, on the other hand, are specifically designed to explain certain ML models. Two of the most widely used post-hoc explainability techniques are namely LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (Shapley Additive Explanations). LIME is a model agnostic technique based on the creation of local surrogate models that are capable of accurately explaining the predictions of any classifier or regressor. In other words, a local model can be created to effectively approximate the explanation of a black box decision. Therefore, useful information about the behaviour of the underlying model can be extracted by perturbing the input data and observing how the prediction changes. Numerous FIs have adopted this technique to identify the most significant features or variables in the scoring process. For instance, a transfer could be refused due to possible fraud based on the amount of money transferred, the country to which it is transferred and the average monthly expenditure of the person making it (Davenport, 2018). However, LIME's lack of model generalization may represent an obstacle to its widespread adoption in different applications. Similarly to LIME, SHAP is a local surrogate model that is influenced by the concept of Shapley values in game theory. This technique is used to determine the

marginal contribution of each input variable on the output. Shapley values ensure that the marginal effect of a given feature is perfectly distributed across the feature value of the instance (PwC, 2018). SHAP can be used to explain how the default probability of a client was formulated.

Conclusion

As ML algorithms are increasingly being adopted in critical applications across the financial sector, the risk of negatively impact businesses and society may escalate if governance or quality assurance is not properly addressed. XAI can fill this gap by ensuring that a machine's decision-making reflects ethical and regulatory guidelines as well as business objectives. Also, if executives are required to accept accountability for their AI systems, they must understand ML's decision-making in order to introduce additional risks to their risk profile. XAI is a human-agent interaction problem which means that humans can ensure compliance by dictating ethical and governance constraints from the very first code line.

However, since modern ML models do not guarantee explainability by design yet, businesses and developers must be equipped with a suite of external XAI techniques. More specifically, current research is focusing on both post-hoc explainability techniques and causal constraints by design. All these XAI methods exploit transparent (logical and linguistically inspired) representations to ensure the interpretability of AI-enabled decision-making overall process. Therefore, in order to foster the development of AI in the financial sector, FIs need to pay particular attention to both symbolic and sub-symbolic techniques that are capable of opening black boxes.

4.2 BIG DATA ANALYTICS IN DIGITAL FINANCE

Financial and insurance institutions can increase their competitiveness through an effective governance of their Big Data Analytics systems. In fact, financial and insurance sector is characterized by very rich information assets which are essential for institutions operating in that sector, and therefore it is crucial to pay close attention to everything related to the entire data processing life cycle. The time invested by data scientists and engineers of financial and insurance institutions on creating Big Data Analytics (BDA) applications (Figure 19) is differently divided in data collection and preparation, analytics development, testing and deployment. All these activities are usually complicated, in most of the cases, by the adoption of a plethora of ad hoc tools, plug-ins, custom scripts, that prevent organizations from simplifying such tasks. In addition to that, the adoption of different technological stacks leads to a high risk of integration issues, high costs and increased complexity.

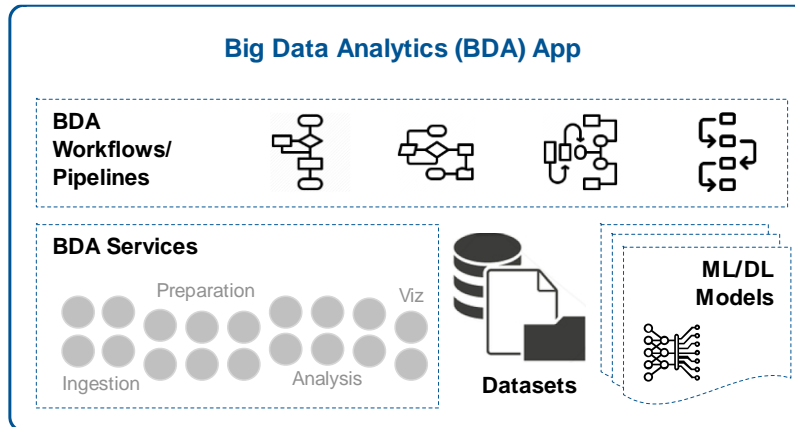


Figure 18 - Big Data Analytics applications

Data Science and Machine Learning (DSML) platforms aim to get an integrated environment supporting products, components, libraries, and frameworks (including proprietary, partner and open source) to design, deploy, execute, and monitor BDA applications, by means of a unique Graphical User Interface (GUI) to guarantee a complete user experience where all components are reasonably interoperable to support an analytics pipeline.

Gartner defines a DSML platform⁵⁹ as a cohesive software application that offers a mixture of basic building blocks essential both for creating many kinds of data science solution and incorporating such solutions into business processes, surrounding infrastructure and products. ML is a popular subset of data science that warrants specific attention when evaluating these platforms. That view from Gartner suggests tackling those issues presented at the beginning thanks to the adoption of a DSML platform based on advanced frameworks able to cover the entire chain: design, deployment, execution, and monitoring of BDA workflows (both stream and batch). Such a platform should be also cloud-native and based on a resource manager solution (e.g., Kubernetes) in order to be able to properly scale computation and storage resources accordingly to the performance requirements of the analytical pipeline.

Focusing on the creation, deployment and execution of pipelines, the ideal platform should be able to process huge amounts of data, in both batch and stream mode, ensuring flexibility and efficiency at the same time. These requirements are implemented at the architecture side by adopting the micro-services approach that proves to be the most suitable solution. That approach has the advantage of automatically distribute each one of those services, thanks to a resource orchestrator (e.g., based on Kubernetes), in a scalable way on multiple servers and infrastructures. To create applications structured as micro-services brings also a resilience benefit when they run in production systems because, thanks to the proper automations, it is possible to isolate the faulty single service and prevent cascading failures that cause the entire app to become unusable. Those micro-services are compliant with development methodologies based on CI / CD, including benefits on the operational side when enabling distributed deployment and continuous monitoring throughout the life cycle of the BDA workflow. Typically, a micro-service contains, in addition to the specific application logic, also all the dependencies required in order to obtain a fully consistent package.

⁵⁹ <https://www.gartner.com/reviews/market/data-science-machine-learning-platforms>

From the developer's interaction side, the DSML for BDA platform should provide a web-based GUI for the design of data pipelines. It has to allow the composition of elementary services for data acquisition and ingestion, data preparation and analysis. Once a pipeline is ready, tested and deployed within a target cloud infrastructure, it can be executed on-demand or according to predefined scheduling logic (in relation to specific processing requirements). That platform should also offer functionalities for monitoring the execution of BDA pipelines in order to properly react and avoid critical situations. Execution, scheduling, and monitoring functions have to be accessible and usable both through the Web interface and through specific APIs.

In order to facilitate the composition of BDA pipelines, such a platform should provide a catalogue of micro-services (the building blocks of the BDA pipeline) that covers all phases, from ingestion to preparation (e.g., filter, merge, cleaning, transformation), to analysis (e.g., machine / deep learning) and finally, publication of results. That catalogue should enable the publication management of the micro-services in terms of additional ones, updates, and deprecation. This kind of interaction with the catalogue should be implemented at two levels: (i) web interface, for human interaction and (ii) via APIs, for integration with the automatic deployment system.

On the data integration side, one of the most important aspects is the availability of multiple data sources (such as ftp, jdbc, mail, mongoDB, mqtt, rabbit, sftp, twitter, websocket) and multiple options of data sinks (such as Cassandra, ftp, hdfs, jdbc, log-kafka, mongoDB, mqtt, rabbit, redis, sftp). The platform has also to provide features for the management of so-called data assets (datasets, trained ML / DL models, processing results, etc.), exploiting the native features offered by the storage layer, supporting the most popular storage frameworks and technologies (such as HDFS, Cassandra, Redis, Presto, Hive, Elastic).

In conclusion, the DSML for BDA platform has to support the full process for creating Big Data applications, taking into account design, development, deployment, execution and monitoring phases. Once the BDA applications are developed, they will then be automatically deployed and installed in their target infrastructures, with the support of package managers that simplify the deployment activities within, for example, cluster architectures.

4.2.1 Trends and Challenges

In recent years, financial organizations are confronted with the challenge of managing very large amounts of structured and unstructured data, which sometimes reach the order to petabytes. The management and analysis of these data can provide exciting applications opportunities ranging from customer centric data analytics that help financial institutions to anticipate the behaviour of their customers and to optimize the services offered to them, to credit risk scoring and innovative data driven techniques for regulatory compliance (e.g., RegTech techniques). Nevertheless, the need to manage such large amount of data introduce several challenges as well including:

- **“Siloed” Data Systems:** Financial organizations possess and manage multiple datastores and databases, including for example operational databases, analytical databases (e.g., OLAP systems and data warehouses), as well as data lakes and Big Data databases. These systems are usually isolated from one another, or in the best case loosely connected. This is a limitation when it comes to intelligently exploiting all the available data in the organization (e.g., towards creating a unified customer profile). In most cases banks and financial organizations spend significant amounts of time and money in data migration from operational systems to data warehouses by means of

Extract Transform Load (ETL) processes. Apart from being expensive the latter have many limitations when it comes to supporting real-time applications.

- **Poor Interoperability:** The integration of data from diverse sources is also challenging due to the lack of semantic interoperability across the data from the different sources. As data arrive from different systems and in different formats, banks have problems combining them and analysing them in seamless and effective ways.
- **Real Time Operations:** Conventional banking infrastructures are developed and engineered for transactional applications that comprise data-at-rest, rather than data-in-motion. Furthermore, analytics were performed following data consolidation that happened overnights. In today's competitive, globalized economy that operates 24x7, there is a need for real time analytics over diverse data that become instantly integrated and instantly available for analysis.
- **Federated identity and access management:** Financial organizations are still challenged by the need to offer customers a seamless experience from access to their data. For instance, customers are still required to login into multiple systems to access and consolidate their Big Data. Therefore, there is a need for a federated identity management and access control infrastructure which facilitates the seamless, yet secure exchange of data across different stakeholders as soon as the proper policies are in place.
- **Regulatory Compliance Challenges:** Recent regulatory developments facilitate data generation and availability. For instance, the PSD II enables financial organizations to access data from other institutions given the consent of their customers, while introducing new value-added actors in the Digital Finance BigData ecosystem. Nevertheless, the explosion of available data makes compliance to other applicable regulations (e.g., GDPR) much more challenging.

4.2.2 Emerging Big Data Solutions and examples of interesting applications

There is a very large number of Digital Finance solutions and applications that are empowered by Big Data or related to them. They can be categorized in different ways such as for example department of the bank/finance they address (e.g., wealth management, investment banking, retail banking, corporate banking). Another broad taxonomy of the various applications can be based on whether they concern the back-office, the middle office, or the front office of the bank. Some the most interesting applications and solutions include:

- **Data Driven Credit Risk Scoring (especially for SMEs):** This application leverages large volumes of data for scoring clients' credit. Instead of relying on the reports of credit bureaus, Big Data enable the use of statistical methods or other data mining methods such as ML. Credit Scoring based on Big Data has the potential to lead to more accurate scoring, especially where diverse sources are used to profile the customer. Most importantly, it is a very important application for scoring SME, which cannot be adequately evaluated based on conventional methods. This is a serious issue, not only for banks and financial institutions, but for the economy in general, as even innovative SMEs with disruptive potential tend to be excluded from access to finance through the banking system.
- **Fast and Intelligent Anti-Money Laundering (AML):** Big Data opens new horizon in AML applications for Digital Finance. The consolidation, integration, and processing of large amounts of data for customers and their transactions can provide AML insights in ways that were hardly possible beyond the advent of Big Data. Specifically, Big Data

management and Big Data analytics technologies enable financial organizations (including regulators) to examine and audit customers interactions at the granularity of individual transactions. Big Data applications enable the flagging of suspicious transactions considering not only the parties involved, the amount and the timing of the transaction, but also its correlation with many other “suspicious” or “fraudulent” indicators. For instance, Big Data analytics enable the identification of potentially suspicious pattern of transactions where the individual transaction might belong. Furthermore, Big Data analytics boosts the identification of hidden patterns of money laundering behaviour, by correlating a given transaction with other historic transactions or other batches of suspicious transactions. Moreover, Big Data facilitate the collection and correlation of linked data from many different systems and data sources beyond conventional banking systems. Such systems and data sources include the Dark Web and cryptocurrencies networks (e.g., Bitcoin) where money laundering and other fraudulent activities take place. Also, Big Data provides the means for accessing and analysing data from alternative data sources such as information about sentiment and reputation of specific people and activities in social media. Alternative data sources boost the accuracy of AML activities, while enabling the discovery of previously unseen patterns of suspicious activities. Finally, Big Data technologies such as the management of high velocity streams can give a significant boost to the speed of AML activities. Speed is in several cases of the essence, as it enables the timely arrest of financial crime and avoids expensive and time-consuming due processes.

- **Intelligent Personal Finance Management (PFM) and Business Finance Management (BPM):** In the scope of the Digital Finance landscape, consumers and businesses are confronted with the task of managing numerous payments and financial interactions with multiple parties and through different channels. State of the art digital finance applications (e.g., mobile banking apps) provide the means for managing these interactions through a single-entry point towards saving time and effort for consumers and businesses. Moreover, they provide a wide range of comprehensive reports. However, the ability of these applications to automate the financial planning and financial management of individuals and small businesses remains quite limited. The advent of Big Data systems and applications can unveil the potential of automated and intelligent financial planning, through providing predictive functionalities that plan ahead in time. The processes of large amounts of historical data and financial interactions (e.g., payment transactions, transfers, credit card transactions, investments, tax obligations, utility bills) for individual and businesses can give rise to the credible planning of future financial obligations for consumers and businesses. This predictive planning enables the next generation of PFM and BFM applications. The latter can greatly benefit from increased data availability as a result of PSD2 and Open Banking. Specifically, PSD2 enables financial organizations to gain access to larger volumes of historical transactions about consumers and businesses, given of course that consumers and businesses consent to this access. In this way, PFM and BFM applications can profile organizations and consumers in a more credible manner towards providing focused and effective financial planning recommendations.
- **Personalized Wealth Management, Asset management and Investment Recommendations:** High quality investing and asset management services are typically assessable to wealthy citizens given their high fees and costs associated with them. Asset managers and the wealth departments of banks prioritize rich customers when providing their investment recommendations and related asset management services. With the advent of Big Data technologies however, they are provided with opportunities for automating a significant part of their asset management advice. This provides a foundation for lowering their costs and providing their services to a much

larger pool of customers. Specifically, leveraging on customers' profiles, Big Data systems and applications can provide automated, personalized, yet credible investment recommendations to customers. In the medium and long terms this can make high-quality investing more accessible to consumers and businesses i.e., Big Data can contribute to democratizing investments and asset management. Note however that any Big Data systems for personalized investment recommendations must comply with applicable regulations and directives like MIFID II (i.e., the Second Markets in Financial Instruments Directive)

- **Robo-advisors:** Similar to personalized asset management systems, robo-advisors automate parts of the investment management process, while sometimes providing trading functionalities as well. Therefore they can benefit from the use of Big Data and analytics (including Machine Learning) for customer profiling and the provision of recommendations about portfolio construction and management. Moreover, they can also leverage Big Data technologies for the management of real-time in cases like high-frequency trading.
- **Usage Based Insurance:** Usage based insurance is one of the most prominent examples of IoT enabled Big Data Applications. It is based on the acquisition of IoT data and their use for the accurate calculation of insurance premiums and the launch of novel usage-based insurance models (i.e., Insurance as a Service). Insurance as a Service models can be implemented in areas like healthcare insurance (e.g., insured individuals wearing fitbits and other fitness devices), car insurance (e.g., using data from the drivers' behaviour), as well as home and property insurance (e.g., using sensors to assess the status of the property and calculate premiums accordingly).
- **Intelligent Market Surveillance:** Big Data systems enable improvements in capital market surveillance towards automatically identifying market manipulation efforts and patterns that deviate from regulatory rules. The application bears similarities to the AML and fraud detection applications. The use of large volumes of data (including alternative data) provides accuracy, along with the possibility to capture market manipulation efforts that would otherwise go unwatched.
- **Regulatory Reporting and Compliance:** Big Data analytics can be also used for regulatory compliance purposes, as part of the wave of RegTech applications. In this direction, large amounts of data can be analysed to identify possible regulatory violations and to audit compliance to applicable regulations. The simplest possible example concerns GDPR auditing based on the audit trails and logging over databases.
- **Personalized Customer Centric Retail Services:** The integration and consolidation of customer data from diverse data sources including operational data (e.g., payments), alternative data (e.g., social media and internet feeds) and data from other financial organizations (e.g., PSD2 derived data) provides the means for more accurate profiling of customers. This can subsequently facilitate the creation of new customer centric products, as well as more effective processes like KYC and customer centric analytics (e.g., in terms of the identification of loyal customers or high value customers).
- **Big Data Databases:** At their infrastructure level, financial organizations can greatly benefit from Big Data databases i.e., integrated databases that bring together all different types of data including static data and data-at-rest, but also operational and analytical data. Big Data databases can revolutionize the ways financial institutions manage their data through saving them time and costs from the currently cumbersome ETL processes. Big Data databases will typically automate and accelerate the ETL processes (i.e., real-time ETL), which decouples financial institutions from the need to migrate operational data to analytical databases over nights. Overall, Big Data

databases unify the management of financial data, while enabling the execution of analytics over “fresh” data rather than on data that are synchronized in daily or even weekly timescales. As such they are also a foundation for real-time analytics.

- **Regulatory Sandboxes:** Big Data technologies are among the main elements of regulatory sandboxes for digital finance, given that most sandboxes are destined to support and facilitate experimentation with data-intensive applications.
- **Finance Sector Data Spaces:** In conjunction with Industrial Data Spaces technologies, Big Data technologies can be used to establish data spaces for digital finance, which will facilitate Digital Finance stakeholders to exchange data in interoperable, secure, and trusted ways. In this direction, semantic interoperability techniques and related ontologies for the finance sector (e.g., FIBO) can be used. Data spaces can facilitate many of the above listed applications, especially in cases of processes (e.g., credit risk assessment, customer profiling) that can be performed jointly by more than one financial organizations (i.e., participants in the data space).

5 RECOMMENDATIONS TOWARDS FUTURE RESEARCH & INNOVATION AGENDAS

This White Paper provides a broad, even though not exhaustive, view of topics, technologies, and applications that are innovating the financial sector as well as enhancing future developments aligned with the evolving market needs. One of the main purposes of the White Paper is indeed to gather several viewpoints appertaining to different stakeholders, as well as different market sectors, to report effective and useful information along with some tailored and straightforward recommendations. Based on the Vision of the White Paper, those recommendations are aimed to set the basis for understanding current market trends, the impact on business aspects, and the shaping of relevant business scenarios. Moreover, apart from highlighting the utilization and exploitation of emerging technologies, a main objective is to present and share suggestions on problems' solutions, using the current state of the art and previous tests as a basis for the experience. The main sources of the information provided in this section are the result of stakeholders' consultations and the collection of the selected use cases (Annex II).

The following table summarizes different challenges and recommendations highlighted by Technological providers while developing solutions, mainly AI-based, to Financial Institutions Clients.

Challenges	Recommendations
Gathering enough data to efficiently train AI algorithms	Transparency and proper communications to foresee in advance the effort expected for training
Limited access to Banking Institution historical data because of regulations	Small scale projects with interested financial institutions and a continuous cooperation among stakeholders
The expected learning curve for the ETL process	Technological choices can allow for quick design and implementation processes, while ML/AI approaches can enable some degree of automation in the initial exploratory stage
Compliance to standards for the PD modelling process	Any solution has to meet the requirements set for model validation, while regular model validation ensures monitoring of performance and stability
Challenges related to the market adoption of a new solution, i.e., platform with a network of companies in direct communication with credits provided by banks	Simplify to the maximum possible extent the onboarding and management processes, in order to offer customers a user-friendly interface, and reduce to a minimum the friction that may arise when banks adopt new systems and workflows

Even though this White Paper does not deeply cover Blockchain technologies and applications, some selected use cases are based on those technologies. The table below represents some of the current challenges related to DLT technologies as well as some possible solutions or recommendations.

Challenges	Recommendations
Blockchain can fundamentally change existing business processes, and	Focused use cases and close integration with existing processes, that together converge into a

through doing so replace legacy systems and ways of operating	more comprehensive shift over time, whilst avoiding a 'rip and replace' type solution
Blockchain technology has not yet reached scalability maturity	Technological developments are improving the scalability, and decentralized and more scalable solutions are now to be selected a preferred in respect to less scalable solutions

From a study made by Accenture in 2021⁶⁰, organisations are redistributing the IT budget and expenses, shifting from a 30-40% to a 60% in Innovation, from 2019 to 2021 and beyond. Moreover, the overall rate of adoption of technologies is increasing consistently, from a 75% to a 95% in AI & Automation, the study reports. Such trends show and highlight the increasingly importance of Innovation and more specifically of the utilization and exploitation of technologies such as AI. The financial sector follows the same trend, however, it faces specific challenges that are hereby reported: (i) data fragmentation; (ii) data availability barriers; (iii) no validated business models; (iv) impact of new technologies on banking business models; (v) sector-specific regulatory barriers. Such challenges are described in Section 0 of this White Paper and shall be thoroughly taken into account while scouting new possible offerings, solutions, or innovations to be implemented in a company's business.

The financial services global market is increasing steadily, and it is expected to reach \$35.000 B at the end of 2030, compared to slightly more than \$20.000 B in 2019. The fintech market, following a similar trend, is expected to value \$325 B⁶¹ at the end of 2030, accounting for almost 1% of the overall financial sector market. As a market in expansion, it opens several opportunities that shall be exploited by bringing innovation and disruption. However, security of consumer data is a major factor that could hinder the growth of the fintech market, and as such it shall be highly considered. Another fundamental trend that must be highlighted, is the growth and the expected size of the fintech market by type of service.

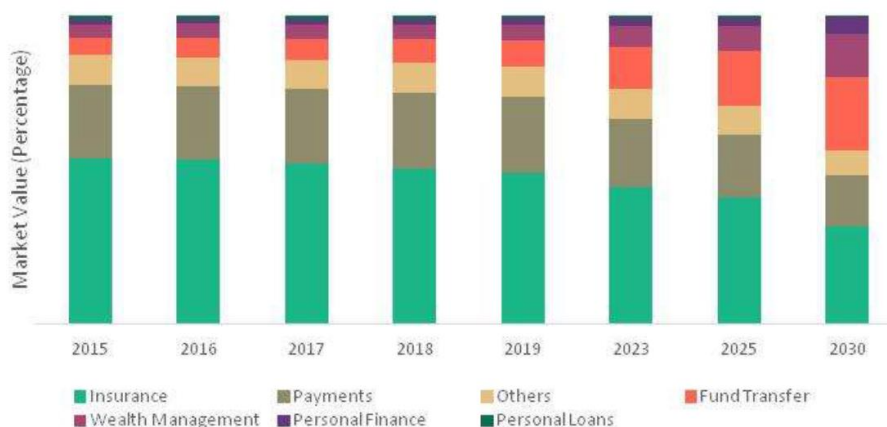


Figure 19 – Fintech trends and market value by type of service, source: *TheBusinessResearchCompany 2021*

Clearly, the fund transfer market is expected to be the fastest-growing segment and it shall be considered for developing and implementing new innovative solutions in fintech sector.

60 https://www.accenture.com/_acnmedia/PDF-153/Accenture-Make-The-Leap-Take-The-Lead-Report.pdf

61 <https://www.thebusinessresearchcompany.com/report/financial-services-global-market-report-2020-30-covid-19-impact-and-recovery>

Advance analytics, AI & ML are among the most widely adopted and foreseen to be adopted technologies in the financial sector overall, and are considered to be key factors for innovation. Moreover, data-driven decisions are taking the lead in any decision making methodology adopted by companies, and matching the right technologies, with the right service, with the right business model, may enhance the chances of penetrating and disrupting the market. In order to do that, a comprehensive understanding and useful utilization of data is needed, indeed 89% of financial services organisations are working to improve data literacy, according to a Exasol's survey⁶².

The provided information are examples, extracts, and analysis of much more detailed data that can be found in the paper as well as in the final Annexes. The importance of such recommendations lays in the approach expressed by the challenges and the ways to overcome them. Lastly, this chapter has been drafted starting from the performed activities during the development of the White Paper, whereas its message has to be broader and indeed will be enhanced in the second version of the paper by making a step forward to the current status and by gathering and researching various and dense data to provide valuable feedbacks and suggestions to the readers.

62 <https://www.exasol.com/resource/whitepaper-financial-services-how-to-accelerate-data-driven-decision-making/>

ANNEX I - DEEP DIVES ON PROJECTS & APPLICATIONS

I.1 DIGITAL NATIVE BANKS

Considering the Fintech and Banking state of the art, technological evolution and growing predisposition of users to approach banking services almost completely through the digital channel have led the spread of new digital native banking entities. According to a recent research⁶³ developed by LINKS Foundation in collaboration with the University of Turin, these entities can be defined as Digital Native Banks. The term Digital Native Bank(s) considers recently established organizations (the phenomenon becomes significant starting from 2013) that offer banking services mainly through the digital channel, supported by a lean technological architecture (without the burden of legacy systems) specifically designed for exploit the latest innovations on data management, in order to offer a superior user experience. Although the term Digital Native Bank is clear and proper to identify these entities, other terms like Neobanks and Challenger Banks are much more widespread despite sometimes the two terms are used indifferently as synonyms, other times they are used to identify deeply different entities. The lack of a common language generates confusion between practitioners and academics, and does not allow the clear diffusion of specific knowledge.

In the developed typology, LINKS Foundation has identified five types of Digital Native Banks (Beta Banks; Neo-banks; Challengers Banks; Big-Tech's Banks; Retailer's Banks) considering five dimensions of analysis (License; Actors; Approach; Banking Market Experience; Group Core Business). The purpose of the typology is to promote a better and more specific identification of the various entities operating in the digital banking services market, encouraging a better understanding of the ecosystem through the diffusion of a common language. The following description aims to reconstruct a comprehensive map of the existing types of Digital Native Banks and their main characteristics.

63 Buchi, G., Fasolo, L., Cugno, M., Zerbetto, A. & Castagnoli, R. (2019). New Banks in the 4th Industrial Revolution: A Review and Typology. In *22nd Excellence in Services International Conference*, Università di Verona

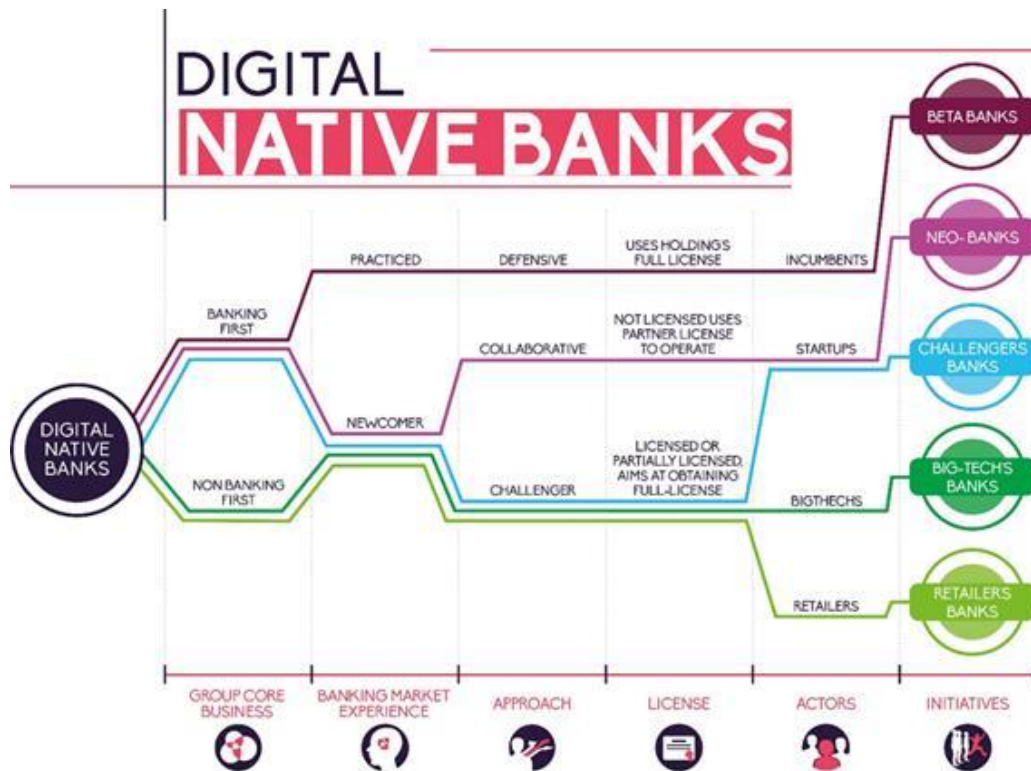


Figure 20 - Digital Native Banks

Approaching the Digital Native Banks landscape, the first analysed dimension for the typology definition is the “Core Business” of the business group in which an entity belongs. It is possible to distinguish Digital Native Banks between two parts: entities belonging to corporate groups mainly focused on banking industry (Banking First) and entities belonging to corporate groups mainly focused on other industrial sectors, such as technological or retail sectors (Non-Banking First). Differently from Banking First, entities from Non-Banking First can have access to resources and skills outside those of the financial sector. A further dimension of analysis concerns the “Experience” within the banking sector, that differentiate spin-off entities of incumbent banks (Practiced) from Newcomers, representing entities that face the market for the first time. Practiced entities can take advantage of a deep knowledge in banking stuff, a consolidated customer base and a favoured access to capital. On the other side, newcomers have to deal with both regulatory complexity and efforts to intercept disaffected customers of existing banks. The different business strategy of actors is summarized by the “Approach” dimension, expressing how a Digital Native Bank relates itself with the *status-quo* of the market. While some new entrants hold a challenging approach against existing banks, by leveraging their digital assets, other entities have adopted a collaborative overture, looking for synergies and reducing execution risks (in particular regulatory risks and the lack of an established customer base). The “Defensive” approach is the one assumed by incumbent banks while setting up their digital native spinoff, considering that these initiatives aim to preserve their market share and stem the spillage of customers. Another dimension of analysis is the ownership of Banking License, that represents a legal prerequisite in order to undertake the banking business. There are three main approaches used by Digital Native Banks in order to operate under the proper banking license:

- entities belonging to banking group often use the parent bank license;

- other entities use the license of a partner bank, often in exchange of commissions;
- other entities have undertaken the path to obtain their own banking licence. For these subjects, given the complexity of obtaining bank licenses (especially in terms of compliance), a modular licensing strategy is structured (e-money or e-payment).

Due to the high pressure from different actors approaching this segment with their initiative, the Digital Native Bank environment is becoming very competitive. Among these actors, in addition to traditional banks, there are completely independent entities that are born as start-ups, large technology companies (for the moment only in China), as well as some groups operating in the retail sector. From the analysis of the previous mentioned dimensions, the LINKS Foundation's typology for Digital Native Banks identifies 5 types of entities: Beta Banks, Neo-Banks, Challengers Banks, Big-Tech's Banks and Retailer's Banks.

"Beta Banks" are new spin-off organizations of traditional banks or joint ventures in which traditional banks have corporate control. By leveraging the parent company's full license, these entities are able to offer almost all financial services. In addition, these organizations enjoy a consolidated experience in the banking sector and a large base of potential customers, although they mainly focus on the millennial segment. From a market point of view, these entities represent a defensive reaction of incumbents from other digital attackers.

"Neo-banks" are Fintech independent start-ups that approach the market as new entrants. Despite they usually do not own their banking license, they have banking as their core business. Often, these entities sign strategic partnership with a fully licensed bank to expand the range of retail banking services offered. This collaborative approach enabled a lot of advantages, such as entering in the value chain of incumbents, focusing on fulfilling a core set of customer needs and getting a strategic partner with strong capitalisation and experience.

"Challengers Banks" are operators that own a full banking license, being able to offer a wide range of banking services. These new entrants want to compete by challenging the consolidated players directly. Mainly based in the United Kingdom, where Fintech trends has grown more than other countries, they benefit from a new, young, dynamic, and digital image which identifies them as the subjects that are disrupting the market.

"Big-Tech's Banks" are organizations formed by large technology companies that do not have the banking sector as their core business and are not experienced in this sector. As new entrants in the banking sector, they intend to challenge the *status quo* leveraging their technological assets and digital capabilities to provide a superior user experience. In addition, they can exploit the trust of the group's brand. Although the main examples of these organizations are confined to China, it seems possible that this will also happen in the West soon.

Finally, "Retailer's Banks" are organizations formed by large Retail groups, which do not have the banking sector as their core business. In addition to the attractiveness of banking activities, these groups see the opportunity both to finance customers to encourage the purchase of the group's core products or services, and to obtain additional information to better know their customer. By creating their banking platform, these groups aim to challenge the status quo leveraging the trust of their group's brand and their distributed physical presence as an asset, despite their lack of banking experience.

I.2 DECENTRALIZED FINANCE AND COMPARATIVE ANALYSIS OF EMERGING CROSS CHAIN SERVICES

Decentralized Finance (DeFi) represents the ecosystem composed by all the applications, developed on top of blockchain infrastructures, that provide access to financial services. In this context, blockchain technology grants a transparent and trustless framework, allowing permissionless access to a variety of financial services, provided that an Internet connection is available.

The decentralized nature of DeFi provides a unique solution to solve critical aspects of the traditional financial paradigm. Firstly, decentralisation eliminates the necessity of trusted third parties, i.e., intermediaries disappear. Secondly, transparency is granted, since all users have access to data concerning all transactions carried out on the blockchains while maintaining privacy. Thirdly, DeFi is able to leverage blockchain technology to foster financial inclusion.

For the DeFi ecosystem to exist, there must be a circulating medium of exchange which in the traditional system we define as *currency*, while in the DeFi context we call *cryptocurrency*. If, on the one hand, fiat money is generally under the monopolistic control of Central Banks, on the other hand, cryptocurrencies represent a form of unregulated digital money that is consensually accepted by the community members of the blockchain.

As mentioned above, DeFi applications (dApps) take advantage of the blockchain technology implemented. Nevertheless, the implementation of most financial services needs the execution of smart contracts⁶⁴, firstly developed on the Ethereum blockchain, that automatically trigger self-enforcing actions arising from an agreement between two or more parties.

The transition from the traditional financial paradigm to DeFi materializes in the coexistence of the two systems which forms a *continuum* between centralization and decentralization in the provision of financial services. The resulting environment is composed by a set of actors characterized by different management systems that make their organization more or less decentralized. In particular, within DeFi, the degree of decentralization generates different implications regarding the functioning and performance of dApps and blockchain infrastructures.

Figure 22 shows the actual DeFi ecosystem from a cross-chain perspective, taking into consideration five categories of financial services: *borrowing and lending*, *exchange*, *deposit/asset management*, *derivatives* and *stablecoin issuance*. In this framework, the Bitcoin blockchain is the only one, among the eight platforms analysed, that does not include the implementation of smart contracts and, in turn, the development of dApps. However, given the impact of the Bitcoin blockchain in terms of network effects, which caused its considerable appreciation since 2012, it represents one of the cornerstones of the DeFi ecosystem. Moreover, the situation outlined by the figure below shows the monopolist role played by the Ethereum blockchain inside the DeFi environment. Indeed, as in the case of Bitcoin, the Ethereum blockchain has generated strong network effects, being the first blockchain infrastructure to implement smart contracts and develop dApps.

⁶⁴ For more details about smart contracts definition, see also [\[Link\]](#).
































































										
	BITCOIN	ETHEREUM	TRON	STELLAR	EOS	TEZOS	NEO	CARDANO		
Borrowing & Lending		    			   					
Exchange / Trading		         	  	   						
Deposit / Asset Management		          								
Derivatives		   								
Stablecoin issuance		      								

Figure 21 - DeFi ecosystem survey across eight blockchain infrastructures.

Nevertheless, more recent infrastructures (e.g., Eos and Tezos) have started to expand their network both in terms of the number of on-chain dApps as well as in terms of financial services offered on their own blockchain. Consequently, the effects deriving from the emergence of other blockchains within the DeFi environment are twofold. First of all, emerging platforms can attract on chain users of other infrastructures, offering higher platform performances to face increasing scalability requirements. Secondly, a richer DeFi environment composed by a multitude of blockchain platforms can bring to an expansion of the decentralised network at the expense of the Centralized Finance (CeFi) ecosystem.

The figure above shows how *borrowing & lending*, *exchange* and *asset management* are the fields within which the DeFi ecosystem has evolved more, driven by the path dependency linked to the first mover advantage of the Ethereum blockchain. However, the development of the environment is moving towards new opportunities not only in terms of infrastructures but also with respect to financial services.

On the other hand, DeFi derivatives define an emerging field in this growing ecosystem and represent synthetic tokens able to reproduce the fluctuations of the underlying assets (e.g., cryptocurrencies, physical assets, gold). In addition, as usual in the DeFi context, these particular tools provide access to a series of advantage, such as tradability (i.e., no requirements to commit to entire high-value investments at once⁶⁵), transparency and autonomy (i.e., non-custodial management of assets). In conclusion, another interesting set of expanding financial tools within DeFi is composed by stablecoins which represent a peculiar form of tokens characterized by limited fluctuation rates. In particular, they can be pegged either to fiat currencies (e.g., US Dollar) or to digital assets (e.g., USDC, TUSD), creating an important incentive in moving also simple payment transactions on top of the DeFi ecosystem.

⁶⁵ In most cases, in fact, transactions conducted on blockchain platforms can involve purchases and sales of portions of assets. For instance, the smallest unit of Bitcoin tradable on the market is called a *satoshi* and corresponds to the one-hundred-millionth part (100.000.000) of a Bitcoin, i.e., 0.00000001 BTC.

I.3 DECENTRALIZED APPROACHES TO MANAGE DATASETS USED BY AI AND BIG DATA SYSTEMS

Decentralized approaches, based on consensus, smart contracts and DLT, may be applied for data governance and management, in AI and ML as «trust» minimizing tools. Consensus can be used to regulate the access, the management and the economic exploitation of datasets in all contexts.

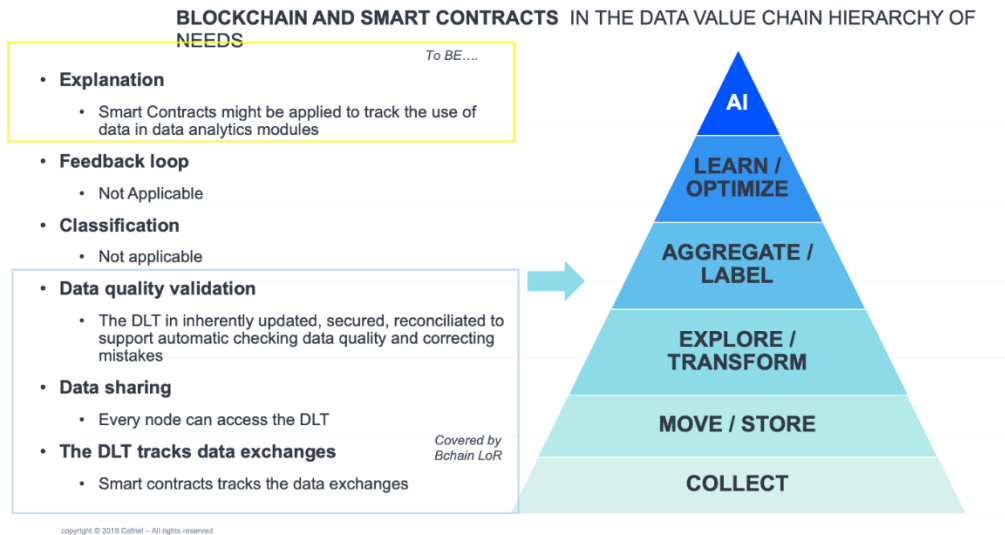


Figure 22 - Data value chain hierarchy of needs

Consensus can help the regulation of use of datasets in many context (privacy, industrial ecosystem).

Blockchain and Smart Contracts are a “full stack” that might be applied to the Data value chain to add transactional logic to the exchange of certain digital asset with a value track the use of data in data analytics modules:

- Tracking and tracings events;
- Automated the execution of logics among peer-to-peer networks;
- Generation of new economic approaches.

Due to the nature of the technology, it brings inherently

- Security of the Data;
- Tracking of any modification;
- Automatic reconciliation.

Blockchain and DLT can be applied to manage the datasets produced by AI systems and model, in three main different ways:

- The dataset can be described via a metadata model and registered on the ledger (notary). Any modification of the dataset can be then registered in the ledger;
- The dataset can be encrypted and the access to the dataset can be managed via a smart contract;
- The dataset can be monetized via a smart contract, and a token approach can be used to define the incentive model of the adoption.

Potential use cases for this approach are:

- Control the access and right to use a certain dataset;
- Enable the exchange between different actors of an ecosystem of certified and tracked data;
- Enable transactions;
- Enable datasets marketplaces.

I.4 NETWORK FINANCE. A SYSTEMIC AND NETWORK APPROACH DESIGNED TO EASE LIQUIDATION OF TRADE RECEIVABLES, OPTIMISE SHORT-TERM CREDIT ALLOCATION, AND ENHANCE AND IMPROVE THE EFFICIENCY OF RISK ANALYSIS MODELS

Network analysis is a specific branch of network science focusing on the study of complex networks.

Studying them implies the use of tools and methodologies borrowed from a variety of areas of knowledge such as graph theory (mathematics), statistical mechanics (physics), inferential modelling (statistics), data mining and data visualisation (computer science), and even elements of sociology.

We're therefore dealing with a realm that is not easy to define or to position. A correct definition comes from the US National Research Council, which defines network science as "the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena".

This instrument has a very wide range of useful applications in several fields: from the spread of diseases to the relations between animals and organisms of particular ecosystems if we consider the biological area, to networks of friendships or relations between members of online communities, if we consider the social area.

With reference to the purpose of this document we'll take into account the main applications in the economic area, and we'll focus especially on the use of network analysis relating to the study of financial relations.

In this case the nodes of a network could represent countries if one wants to study import-export relations between these nations or, as in the case we're about to examine, nodes could represent firms if one intends to analyse economic and financial relations between them.

So, imagine a network where each node is represented by a firm, the outgoing edges (from the nodes) are represented by payments, and the ingoing edges (to the nodes), by collections.

Within such a configuration the study of the network makes it possible to spot the most important or most peripheral nodes in the network; they can make possible to reason on complex phenomena such as the location of productive relations, the efficiency of the allocation of capital, and predictions about the possible spread of financial crises.

In the case of financial networks where links between nodes are represented by credit/debt relations, an accurate study of the degree of interdependence between its nodes, of their centrality or peripherality makes it possible to think in a different way about credit risk and about the propagation of financial crisis, shifting the attention from *too big to fail* to *too*

interconnected to fail.

New instruments to measure the relations between nodes/firms, such as the degree of centrality in its various declinations, the participation to communities that can be reconstructed with the analysis of small-world networks and of the six degrees of separation, or the clustering analysis that study the ways in which firms tend to group together, create different perspectives to look at financial relations, suggesting the creation of new tools aimed at dealing in an innovative way with structural problems in the economic and financial management of banks and companies: from liquidity management, to liquidation of working capital, to risk management, to optimal credit allocation.

Network analysis applied to economic relations could be regarded as a basis for a new approach to financial management in the realm of complex economic systems; this novel approach takes the name of *network finance*.

Let's try now to explain in more detail how this new instrument has allowed, and will allow in the near future, the development of novel financial instruments which will go alongside the traditional ones, offering more efficient solutions to entrenched problems that have so far met only partial solutions.

But how can an approach tied to the concept of network finance foster the birth of new solutions?

Let's take as an example the need, repeatedly expressed by the market, for more effective tools for the liquidation of commercial credits⁶⁶ capable of releasing the liquidity trapped in firms' balance sheets, also because of norms linked to ratings, which make it difficult for MSMEs to access credit.

This kind of problems are well known. In the European Union market, around 40% of sales are made on credit. Italy this figure rises to 42%, with an average of 98 days before payments are made.

Over 30% of Italian firms states that sales made on credit and overdue payments compress the businesses' liquidity, thus creating non-negligible problems.

Losses due to insolvencies have reached 3% of the total. About 25% of all bank loans are made to manage late payments.

Yet two SMEs out of three still lament difficulties to access credit and scarce availability of effective tools for managing working capital⁶⁷.

This already complex picture is bound to get worse because of the economic fallout from the present Covid-19 sanitary emergency.

As everybody knows, the traditional approach to finance regards payments as single relations between one creditor and one debtor.

If instead we frame those disconnected bilateral relations in a multilateral network

⁶⁶ This hypothesis can be placed in the following scenario: the global market for liquidation of receivables and payables is worth more than 16,500 billion euros, 4,000 of which are European, with Italy being the "largest" market in the EU, worth 463 billion (elaboration Polimi on data from Orbis, ECB, OECD, Eurostat, interviews with experts, April 2020). Traditional methods of liquidation and disinvestment, of advances on invoices and factoring systems, even if we take into account native digital systems such as invoice trading and dynamic discount, are able to manage only less than 25% of this outstanding (countries with a different DSO are in a similar situation). This means that a large slice of receivables are left unmanaged and at risk of being liquidated after they come due (Polimi Osservatorio Supply Chain Finance, April 2020). Moreover, solutions currently used are mainly addressed to medium-sized businesses enjoying a high rating, and these solutions have high barriers to entry both in terms of size and of cost and cut off some types of firms. So most businesses choose to wait and just hope that customers pay off their debts.

Even the utilization of supply chain finance techniques, although a step in the right direction, can't get an optimal short-term allocation of capital, being subject to very strict rating evaluations from the banks funding liquidation.

⁶⁷ Source of data: Atradius, Osservatorio Polimi, Intrum Justitia (2019).

perspective, they can be reorganized in a different way, that would prompt all the existing interconnections between creditors and debtors to surface, and thus greatly reduce payment times and deeply modify risk dynamics.

Through the network finance approach, the economic system represented above can be analysed with intelligent algorithms which fathom Big Data, made up of relations not only between nodes but also between communities of nodes. This way, payment cycles capable of self-sustaining with no need for external funding can be spotted.

In concrete terms, each node can be enabled to pay part of their debts while at the same time collecting part of their credits, without the need for specific financing for liquidation of working capital, by using cycle analysis and cycle balancing techniques.

Reviewing this type of relations in a network perspective, it is possible to design tools capable of strongly reducing the time of the cash-to-cash cycle, enabling firms to settle part of the invoices without needing to resort to credit: in short, it is possible to pay part of one's debts with simultaneous collection of one's credits.

The forthcoming launch on the market of fully digitized native *network finance* tools for the liquidation of receivables will make it possible, thanks to a sequence of algorithms, to successfully process thousands of invoices, to locate and analyse millions of cycles and to settle payments by interconnecting positions, thus optimizing the debt burden and improving output in terms of settled volumes.

By viewing an economic system as a network, it becomes feasible to use models that generate equilibrium, involving at the same time many possible entities and freeing the liquidity trapped in the system itself.

Thanks to the information about the network to which a payment belongs, and then studying all the possible combinations that a given payment would generate in the system, the optimal combination of payments can be found. This combination includes all payments that better correspond to the best functioning of single nodes and maximize the financial position of the whole network, and is the foundation upon which tools of targeted credit can be offered, while reducing risk for the entire system.

ANNEX II COLLECTION OF USE CASES

II.1 BANKING SECTOR

II.1.1 Intelligent Document Processing – IDP

Responsible:	GFT
Sector:	Banking
Typology:	Industrial
State of the Art:	On the market
Keywords:	AI, Cost reduction, Process optimization

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Through globalisation and technology innovation, businesses have expanded their reach and their customer bases far beyond the boundaries of their initial immediate locations. With exponentially growing customer bases and increased auditing and reporting standards, highly regulated companies in the financial sector are facing exploding volumes of documents to be processed. These large volumes of semi-structured and unstructured documents have to be analysed, classified and processed with great accuracy and speed.

Most companies are classifying and extracting data from documents and forms manually, which is a slow, cumbersome and costly process. We can observe an increase in the adoption of complementary technology solutions such as **RPA (Robotic Process Automation)**. Most companies are applying **OCR (Optical Character Recognition)** software to do so but this option requires high-personalization and manual configuration. The rules and workflows for each document and form need to be updated often, which leads to continuous maintenance which is unsustainable.

In addition, banks and insurers are facing a fierce competition concerning the efficiency of their processes, the cost and the time needed to treat demands, claims or administrative tasks.

This leads to a growing demand among enterprises to enable end-to-end process automation with integrated RPA and Intelligent Document Processing (IDP) capabilities. The ability to answer customer's needs fast and accurately coupled with a higher degree of automation is driving the finance sector's appetite for highly sustainable and automated solutions.

2. Methodology

Description of the use case and its methodology

The intelligent document processing solution is leveraging Natural Language Processing and Deep Learning in order to classify documents and extract relevant information automatically.

Documents (such as PDFs, Payment slips, claims, law suits or emails for instance) are extracted from **ERPs (Enterprise Resource Planning)**, **CRMs (Customer relationship management)** or Core banking platforms and sent to a document basket. The solution then picks up the documents and first conducts a set of operations to enhance the image. Auto cropping and noise reduction are common elements of this step.

The document is then analysed via OCR and computer vision in order to extract data and tags that will serve as the basis for the machine learning algorithm. Once scanned, the document is handed over to the classification algorithm and then assigned with a label and a confidence level (% of confidence in the document classification). After training, the deep learning algorithm is capable of extracting elements from the documents, such as names, **IBANs (International Bank Account Number)**, addresses and other common elements the human agents would normally identify and manually input in their systems.

Once extraction is complete from the classified documents, the structured data and the produced report await to be processed to their assigned destination, which can also be their source platform.

In order to facilitate and accelerate the process, the solution comes along a set of tools: the tagging tool, which highlights information to be extracted and accelerates the creation of proper examples to train the algorithm; and the validation engine which provides a clear control dashboard of the documents that were processed to control proper classification and data extraction. The correction of errors or cases with low confidence will further refine the accuracy of the algorithms and improve their accuracy over time.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

Classification and extraction are performed using supervised learning algorithms based on deep neural networks and python language. Datasets involved are unstructured documents that need to be tagged in order to train the algorithm.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The outcome is a nearly fully automated document processing. Human intervention is tuned down to its minimum: tagging the documents for the training of the algorithm, and validation and control of the processed documents.

The usage of an intelligent document processing tool provides benefits in the following segments:

- Process efficiency: Greatly reduces the time spent on document processing and the average handing and response time. It improves accuracy and enables higher scalability of the business. Changes in document layout over time are not going to be dealbreakers anymore and can be easily handled.
- Financials: Cost reduction, higher client satisfaction and exploration of new revenue streams through analysis of the collected data.
- Human capital: Employees with higher availability can be repurposed towards more creative and innovative initiatives.

The usage of the tagging tool offers a 150% additional effort reduction in comparison to a fully manual tagging operation.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

The main challenge in any AI problem is gathering enough training data to train the algorithm efficiently. As such, proper communication and transparency are expected in order to familiarize interested parties concerning the effort expected for the training.

Fine tuning and improvement of the solution algorithm and capabilities require good knowledge of artificial intelligence, especially deep neural networks. As such, the client has either to develop inhouse capabilities or depend on the vendor, which could imply some form of lock-in.

II.1.2 Data Driven Decisions – 3D

Responsible:	Wings ICT
Sector:	Banking
Typology:	Industrial
State of the Art:	On the market
Keywords:	Loans, Probability of Default, Credit Scoring, Marketing

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Determining the Probability of Default is of at most importance for Banking Institutions. According to the approach set by the Basel Committee, a borrower is considered defaulted when (a) the obligor is unlikely to pay in full without out giving up a collateral, if held; and (b) the obligor is more than 90 days past due.

Probability of Default (PD) refers to the likelihood that a default event will occur at a specific time-horizon. It is also the basis for determining *credit scores*, i.e., a numerical expression representing the creditworthiness of an individual, and it is closely connected with metrics like *Loss Given Default* (LGD), denoting the expected loss for a bank, and *Exposure at Default* (EAD), i.e., the gross exposure upon default.

Currently, the aforementioned loan risk metrics are calculated under a strict regulatory framework with mainly static datasets encompassing limited information and based on statistical modelling.

The challenge here was to (a) apply advanced ML/ AI techniques for modelling PD and the accompanying risk metrics adhering to the current regulatory framework by offering transparency (i.e., supporting analysis of the results) and using it as a point of comparison and adjustment, and (b) exploiting big data technologies to incorporate information from large datasets.

2. Methodology

Description of the use case and its methodology

The current version of 3D solution is a solution consisting of two main modules; (a) Loan Risk Management module; i.e., AI-based estimation of PD, assignment of Credit scores, calculation of EAD and LGD and (b) Marketing module; including intelligent derivation of customer segments, estimation of customer lifetime value, churn analytics, product recommendation and others.

It relies on: (a) standard and in-house built/ patented methodologies, based on ML and AI; (b) means for efficient Big Data management and for flexible functionality deployment. The focus was to create advanced modelling pipelines -including feature engineering and selection- and to incorporate both the standard data used in risk management frameworks as well as additional behavioural borrower characteristics in the modelling process.

In fact, the 3D solution takes into account highly granular data, i.e., on borrower behaviour, as opposed to usually limited aggregated data employed in the statistical models. This approach helps realise a risk framework where previously unexploited and undetected patterns are included in the risk modelling process. Lastly, it offers not just a simple implementation of the model, but allows for detailed supporting analysis of the results. This way, it may be used alongside the existing techniques, for comparison and adjustment and/or combining approaches while understanding their limitations and therefore meeting regulatory requirements.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

The dataset in which the use case has been tested is based on information commonly found in a core banking system, but the solution is diverse and may incorporate additional and/or external data. The current dataset includes obligor characteristics, loan characteristics, historical data on defaulted cases and recovery data, going back many years.

The solution itself follows the best practices in data-driven development (DataOps) and incorporates the latest advances in Big Data and AI/ ML. In more detail, it follows a *microservice architecture* and its design allows for the provision of the 3D solution either on premises or on cloud. It can access various diverse *SQL/ no-SQL data sources*, as well as streaming data through publish-subscribe mechanisms. Additionally, ETL pipelines/ workflows can be designed and adjusted to the needs, while data are stored in a distributed file system that allows for streaming access. Furthermore, the solution allows for the use of state-of-the-art frameworks and libraries for the execution of the proprietary ML/ AI pipelines and methodologies, as well as de facto standard technologies in service-based architectures for the orchestration of containerized microservices, automating deployment, scaling, and management.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The outcome of the aforementioned solution is sufficiently validated and is compliant with the Basel validation framework. The 3D solution offers its estimates of **PD, LGD (Loss Given Default) and EAD (Enterprise Application Development)**, using also the statistical baseline approach for comparison reasons. Additionally, statistical tests and combination tests with multiple measures (i.e., the mean differences, maximum deviation, Gini coefficient, and Information Statistic, **ROC (receiver operating characteristic)** curve statistic, and others) that can be collapsed into just one statistic are offered. Furthermore, model validation is performed regularly, so as (a) model performance and stability is ensured and (b) trigger mechanisms may be created to inform risk managers on performance deviations.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

The main challenges phased include:

- The limited access to Banking Institution data; also, under regulations, underlying historical observation period must include many years, whether for external or internal data;
- The expected learning curve for the ETL process; i.e., data from different Banking Institutions and the exploration process of the validity of dataset available;
- Compliance to standards for the PD modelling process (Basel II).

Access to data was achieved through small scale projects with interested financial institutions and a continuous cooperation with a large integrator providing solutions for banks -among others- worldwide. While the expected learning curve for the ETL process cannot be avoided in full, the technological choices made can allow for quick design and implementation processes, while ML/AI approaches can enable some degree of automation in the initial exploratory stage. Lastly, any solution –as is the case for 3D- should meet the requirements set for model validation, while regular model validation ensures monitoring of performance and stability.

II.2 FINANCIAL SERVICES

II.2.1 Blockchain in the Digital Asset Supply Chain

Responsible:	CeADAR
Sector:	Financial Services
Typology:	Industrial (Exertis Supply Chain Services, Sonalake)
State of the Art:	Under construction
Keywords:	Supply Chain, Digital Assets

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Complex modern supply chains, and in particular technology product supply chains, are globally distributed, partially connected networks comprising a diverse collection of individual participants. These include sub-component manufacturers, component manufacturers, finished goods manufacturers, transport & logistics, distributors, retailers, resellers, right up to the end consumer. These entities, whilst co-operating with each other in one-to-one contractual business relationships, are also in many cases competitors, operating with limited trust and only sharing data on a need to know basis. This makes it extremely difficult to manage or implement end-to-end solutions, such as ensuring traceability and provenance, overall efficiency improvements or leveraging advanced analytics and AI.

Moreover, as assets become more digitised (value represented in digital tokens, licence keys, software and digital good etc) their journey and lifecycle through the supply chain, and beyond amongst end consumers, becomes even more challenging to manage.

Blockchain and DLT offer a solution to address these issues in complex supply chains such as these. Data can be shared between participants securely, and in a trusted manner and an end-to-end view of the supply chain can be enabled which supports advanced analytics and AI.

2. Methodology

Description of the use case and its methodology

A key challenge with the adoption of Blockchain and DLT solutions is that they are often designed to radically replace existing business processes and systems. This can be too disruptive for many established industries, especially then there are many individual stakeholders such as in complex supply chains. Our approach in this project is to address key aspects of supply chain processes individually, in an iterative way, thus introducing Blockchain step by step. Once trust is gained in the new technology, and tangible benefits are experienced, the industry becomes more receptive to broader adoption.

The project is initially focussing on integrating a distributed, Blockchain based data model that integrates with legacy systems used across the industry (such as from SAP) to address very specific use cases. The first of these which is near completion is the distribution and sale, and potential re-sale of digital activation keys (a digital asset) in the gaming industry. Whilst using Blockchain alongside existing systems and processes limits some of the full potential that a distributed ledger model could offer, enabling its adoption in this way creates a pathway to full adoption, and all the benefits of a shared, distributed data model.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

The project is being developed on Hyperledger Fabric, an open-source distributed ledger platform that is supported by large and diverse collection of organisations, and is used across a wide range of different applications. As Blockchain is still in the relatively early stages of technology maturity, there are many technical aspects specific to our use case which require research and development activities. In terms of real-world deployment, a number of software service providers are beginning to offer Blockchain-as-a-Service offerings based on Hyperledger Fabric which we are utilising to aid production system deployment.

Current supply chain systems generate large volumes of transactional and other types of data, albeit in a siloed way, and we have designed shared data models based around existing datasets. A key benefit of Blockchain in this space is its ability to support advanced analytics functionality and AI, and we are developing the data models around this to power a number of analytics use cases such as demand prediction and fraud detection.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The Blockchain based solution is designed to bring a broad range of benefits to complex supply chains, including traceability and provenance of physical and digital assets, improved end-to-end visibility and applications of AI and advanced analytics for e.g., improved demand prediction and efficiencies. This is demonstrated firstly through a number of individual, specific use cases and solutions, and then towards a more general paradigm shift in how data is shared and accessed across the supply chain.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

Blockchain is a disruptive technology that can fundamentally change existing business processes, and through doing so replace legacy systems and ways of operating. However, industries are often not ready to adopt such change on a large scale, and this has been the downfall of many solutions as they attempt to move from proof-of-concept to industry adoption. Our project addresses this through iteratively addressing manageable, focussed use cases and close integration with existing processes, that together converge into a more comprehensive shift over time, whilst avoiding a 'rip and replace' type solution.

Technological maturity is another challenge, specifically in following the broader development of blockchain and distributed ledger platform implementations whilst simultaneously building on these to address niche functionality that is not offered through a general purpose platform.

II.2.2 DAOs in Decentralized Finance (DeFi)

Responsible:	Links Foundation
Sector:	Financial Services
Typology:	Industrial (MakerDAO, Swissborg and Uniswap)
State of the Art:	On the market
Keywords:	transparency, financial inclusion, blockchain, decentralized finance, DAO

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

In response to the standard financial industry, Decentralized Finance (DeFi) represents a decentralized ecosystem composed by platforms that are able to provide permissionless access to a variety of financial services, eliminating intermediaries and enabling transparency while ensuring privacy.

Taking advantage of DLTs, Decentralized Autonomous Organizations (DAOs) represent non-hierarchical organizations governed by goal seeking communities united together by purpose and rules that operate through cryptographic routines. DAOs also provide public access to the network and rely on the contribution of the internal stakeholders on a voluntary basis.

In this context, the innovation potential of DAOs within DeFi is twofold: 1. from the technological point of view, DLTs make available an auditable and immutable ledger that contains the history of stakeholders actions and, 2. from a governance and organizational one, DAOs give access to a decentralized environment where incentives are aligned among the different stakeholders and hierarchical arrangements of traditional institutions are replaced by a system of incentives that drives users' interest towards the benefit of the community.

The result is an ecosystem composed by decentralized organizations able to provide access to a wide spectrum of financial services eliminating most of the barriers to entry and thus fostering financial inclusion.

2. Methodology

Description of the use case and its methodology

Within the DeFi ecosystem, each DAO provides access to different financial services. The internal organization structure of every network defines the decentralization degree of the platform. Specifically, depending on the interactions among the actors of a network, it is possible to observe the degree of cooperation which, in turn, defines the level of the network dependency with respect to every actor who interact with the platform. These characteristics of each DAO derive from the application of different governance, which also influence the implementation of the specific financial services.

The table below summarises the main information related to three example of DAO (i.e., MakerDAO, Swissborg, and Uniswap), presenting also the methodology through which decentralization is implemented. MakerDAO and Swissborg present a multiple governance layer, while Uniswap applies a simple but strongly innovative on-chain protocol that makes it the most decentralized DAO among the three. The *decentralized approach*, in turn, identifies the decentralization layer applied in each platform.

	MakerDAO	Swissborg	Uniswap
Financial Service	Borrowing & Lending	Asset Management	ERC20Token Exchange
Governance	Foundation Community	+ Company Community	+ On-Chain Protocol
Role of Community	Voting + Service Usage + Revenue Sharing	Voting + Service Usage + Revenue Extended Workforce	Service Usage + Liquidity Providers
Decentralization Approach	Governance	Organization Governance	& Automated Governance

The information gathered in the table above allows to observe the flexibility behind the application of such organizational approach and its capacity to apply in different context with diverse results.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

The application of a DAO organizational structure also needs the implementation of blockchain technology in order to ensure decentralization in data management. Within the context of the three DeFi examples presented in *point 2*, the utilization of a DAO structure allows to engage the community at an organizational level. Since in these examples DAOs apply in the DeFi context, these organizations also take advantage of smart contracts to trigger self-enforcing actions arising from agreements among two or more parties that generally involve financial services. Particularly, respecting the fundamental blockchain principles of fairness, accessibility, transparency and trust, the application of a DAO organizational structure can apply in many more field also beyond the financial industry.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The three solutions derive a set of important value drivers. Since the three use cases leverage the same technology, these value drivers are common to all the three solutions: *zero rent extraction; cost efficiency; transparency; censorship resistance; ease of use for widespread accessibility*. In turn, applying different governance rules, each use case reaches different outcomes with respect to the financial services provided. The main principle shared by the three use cases concerns the creation of an ecosystem based on trust, fulfilment and inclusiveness. In addition, these three solutions aim to achieve greater alignment between generation and redistribution of wealth between those who generated it and those who internalized it. The expected outcome, deriving from this mission, is the capacity of attracting a broad spectrum of investors ranging from those who do not have access to traditional financial services (i.e., unbanked) to those who are already customers of services provided by CeFi. Moreover, for the nature of the blockchain technology and for the flexibility of DAOs, it can be expected future development of this type of platforms towards the convergence with other emerging technologies such as AI, Big Data and IoT. Till now, Swissborg is an encouraging example of a platform that gives access to financial services based on blockchain technology, employing also AI technology to provide insights about investment strategies.

5. Challenged and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations and lessons learnt

The main challenges that this raising ecosystem is facing include:

- **Scalability:** blockchain technology has not yet reached maturity. In some cases, in fact, it suffers scalability issues that still prevent an even wider diffusion of these solutions around the world;
- **Knowledge:** as in the case of CeFi (if not even more), accessibility to decentralized financial services requires technical knowledge, creating potential barriers to entry;
- **Regulation:** since blockchain is still an emergent paradigm, a number of regulatory aspects still need to be fully cleared. Despite the progress made so far, in fact, some concerns are still present as to how existing regulations may apply to this context.

In this context, technological developments will improve scalability capacity of the platforms. New entrants as, for example, Eos and Tezos blockchains propose more scalable solution with numerous projects within the DeFi environment. Moreover, new entrants, characterized by the presence of a decentralized layer as well as a centralized one, can facilitate the access to knowledge and create bridges between the two ecosystems to meet the regulation requirements of the economic system.

II.2.3 Network Finance

Responsible:	Kitzanos
Sector:	Financial Services
Typology:	Industrial
State of the Art:	Already tested
Keywords:	Settlement, Invoices

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Many companies declare that sales on credit and delays in payments compress the management of corporate liquidity, creating costs and problems that are not negligible in the management of working capital. Losses on receivables account are not an insignificant part of the total. Many SMEs continue to have difficulties in accessing credit (to credit access) and a relevant part of bank loans are allocated to management of late payments.

The challenge is to create on a proprietary platform a network of companies to demobilize their commercial receivables without resorting to credit. A new perspective is to reorganize the single one-to-one relations between debtor and creditor as a network view one to many.

2. Methodology

Description of the use case and its methodology

Financial institutions consider payments among their clients as individual relations between debtor and creditor. Thanks to the new possibilities introduced by PSD2 the company has developed a model and a technological stack to collect and reorganize the classical bilateral information, evaluating the existing interconnections between debtors and lenders according to a network perspective. Without external liquidity and credit risk.

Periodically a proprietary algorithm optimizes the financial flows and the Platform is able to reset to zero the payment time between companies on part of their invoices without credit expositions.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

Based on cloud technology, it offers a fully automated platform. Participating companies define their basic parameters in addition to the list of invoices to be validated in the system. Its algorithms automatically manage the process of payment and management of corporate wallets. The dataset is periodically updated on the relationships, economic and otherwise, between the companies in the network. All this with the help of big data analysis on network models. Regarding the technologies used, the platform consists of different modules connected to each other via the REST API:

1. Platform: a front-end fulfilling current standard, a fully designed back-end developed in Java is based entirely on microservices paradigms, the platform can be easily installed in the cloud or on-premises environments via Docker and/or Kubernetes. The frameworks used for IT

development are all open source and consolidated. Microservices through REST APIs allow easy integration with other systems.

2. Engine: developed entirely in Python, a black box that receives as input a series of information that the proprietary algorithms take charge of to carry out analyses of the relationship networks between companies via Big Data, AI, and ML techniques.

3. Payment system: form for managing funds and payments on the platform managed by a bank or an approved PSP.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The purpose of Bflows is to allow firms and professionals of any size, sector and country, to liquidate their trade receivables without needing to resort to credit. The expected outcome is to enable them to collect their receivables faster and to pay their suppliers more timely without risking insolvency, simplifying and making more efficient the management of and improving their cash flow. Our solution has been tested and fully validated on a dataset provided by one of the 5 largest Italian banks. The KPIs reached are very promising considering the size of the sample dataset. Bflows' algorithms managed to settle 53% of the processed invoices, worth €1.54 B, with an average payment time of 6 days, against the national average of over 100 days.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations and lessons learnt

While the challenge faced by SMEs and professionals is a great difficulty in being granted loans (thus implying further delays on payments and compounding the bad economic situation that made them need a loan in the first place), the challenges we faced in setting up and implementing Bflows have been many-sided.

The first challenge has naturally been a technological one, especially as regards designing and perfecting the core algorithms. In order to make them as mature and effective as possible we resorted to the help of expert mathematicians and physicists; this enabled us to reach the promising KPIs we described in the previous section.

Another challenge we faced has been technological as well as related to market adoption, which implied a twofold effort: on the one hand we had to make the system as simple as possible to integrate with the banks' legacy technology; on the other hand we had to simplify to the maximum possible extent the onboarding and management processes, in order to offer our customers a user-friendly interface, in order to reduce to a minimum the friction that may arise when banks adopt new systems and workflows.

Another challenge we faced is common to any disruptive innovation. The implementation and business aspects of the go-to market phase might have to overcome cultural resistances (such as fear of being disintermediated and difficulty in changing perspective) and problems in modifying operating models (fear of excessive efforts to adopt new tools and changing the status quo). Getting around these hurdles is key in order to reach quickly the critical mass needed to make Bflows a cost-effective tool for the early adopters.

We'll be able to tell whether these implementation and business challenges have been successfully overcome only after the launch on the market of the first network in the Italian market, scheduled for January 2021.

The lesson we learned is that it's very important to start networking from the very beginning of any project in order to be able to rely on partners who would spread the word of and advocate for your product. We would therefore recommend entrepreneurs and innovators to exploit any opportunity of networking and interchange, not least those provided by the dedicated EU bodies.

II.3 INSURANCE SECTOR

II.3.1 Personalized insurance products on IoT vehicles

Responsible:	ATOS
Sector:	Insurance
Typology:	R&I project (Infinittech)
State of the Art:	Under construction
Keywords:	Pay as you drive, Fraud detection, Claim management

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Current insurance policies are based on statistical analysis that exploit drivers' historical data and professional experience to categorize their clients. The datasets and the mechanisms used to classify the drivers are too static and neither include the particularities of each client nor enough flexibility to consider the different situations that happens in a driver's daily journeys. On the contrary, driver's score historically considers age, gender, car's colour etc.

The challenge here is to define and develop new services based on connected cars and AI technologies that support the adaptation of the driver's policy to his/her daily driving, as a modern way of service's consuming fashion: you pay depending on how you drive. This use will clearly result in a more personalized and smarter insurance system, which will also create a basis for automating processes like fraud detection and claim management.

2. Methodology

Description of the use case and its methodology

The reference scenario is a risk assessment service based on real data collection (including IoT) and enhanced by AI and Cloud Computing. Its implementation will develop initially around two business services:

Pay-as-you-drive: a used-based payment service that defines and infers (using different AI/ML, Big Data and IoT platforms) different drivers' profiles based on the way the driver drives and so usage-based prices.

Fraud detection: as a mechanism that analyses real time car data, traffic information and other context datasets to properly depict and relate traffic incidents.

The initial focus is set on Smart Connected Vehicle entities and the datasets collected from these devices. The use case will periodically gather information about the status of a wide number of connected cars and so register historical data of their behaviour. The collected information will be analysed to define and extrapolate different driver's profiles that later assists AI and cloud computing technologies to classify a given driver according to the way he/she drives. This dynamic classification will derive in new adaptive insurance usage-based products. In addition,

data captured from smart vehicles may also feed and enhance roadside assistance and fraud-detection tools, providing relevant context information related to a traffic accident or incident.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

From an overall perspective, this use case will require, as a first component, vehicle smart objects as data sources. These will be provided as two different datasets: as real IoT devices (cars); and simulated smart objects that enhance the AI training processes. Other external traffic data sources will complement the analysed information. The data anonymization building block will guarantee that only traffic and car related information is used to define and train the general drivers' profiles without putting their privacy at risk. Finally, the AI framework will use different AI technologies to, assisted by Cloud Computing and **HPC (High Performance Computing)** components, infer drivers' profiles by processing real time connected cars' collected data. Also, these datasets will assist in possible fraud detection.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The solution will derive a set of driver profiles that map different driving styles by analysing huge datasets from vehicles. Data gathered from a connected car would identify the profile of the current car's driver that, combined with other context information (such weather, location, date and time) can estimate its actual risk (from the insurance company point of view). The number of the different obtained profiles and the matches detecting different drivers (by detecting different profiles driving) would be the initial relevant KPIs.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

The main foreseen challenges would include:

- **Gathering enough and relevant datasets from vehicles that allow the system to define and detect a wide enough set of profiles that cover most of the driver's population**
- **Identify the proper correlations and relevant parameters from the collected datasets that better define and differentiate the profiles and so, the AI models to infer them**
- **- The mapping between the drivers' profiles and the context information to provide accurate risks estimations**

II.3.2 Fraud detection and prevention in car insurances

Responsible:	Rulex
Sector:	Insurance
Typology:	Industrial (GFT, PwC)
State of the Art:	On the market
Keywords:	Predictive analytics, Frauds prevention, Real-time predictions

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Detection of frauds (in particular related to fake accidents between vehicles) is one of the key challenges for insurances. Usually, fraud detection is done based on legacy rules, derived from years of experience in the field. Nevertheless, fraud mechanisms change and therefore legacy rules can quickly become obsolete. For this reason, fraud detection engines should update quickly and automatically starting from the data of the most recent frauds.

In general, any new fraud detection system should be integrated with existing and working systems in order to ensure the continuity and avoid sharp changes of policy. Moreover, explainability of the solution is a crucial topic, too.

2. Methodology

Description of the use case and its methodology

The approach to solve the problem consists of two steps. In the first step existing antifraud rules are automated and applied to any new accident. These rules are directly provided as inputs by expert users. Actually, the goal of this phase is not discovering new fraud mechanisms, but only automating the procedures and testing new frameworks for decision automation in a safe way.

In the second phase new rules are built starting from historical data related to accidents. The generation of new antifraud rules is based on a machine learning approach: based on historical accidents, for which it is known whether they were frauds or not, a supervised classification problem is solved to find rules that distinguish real accidents from frauds.

Accidents reported as frauds by the rules can undergo to further investigations to verify if they were actually frauds: this new record is added to the historical dataset and can be used to build a new set of rules.

Classification methods can also produce a confidence score related to the probability that the prediction is true. In this way, potential frauds can be prioritized in order to avoid overloading assessors with more frauds to investigate than possible

3. Technological Components

High-level description of the utilized technologies and the datasets involved

ML: it includes techniques for deriving directly from historical data without the need of setting any kind of a priori hypotheses. In particular, classification approaches are used: in this case historical data contains a label, which is the output of the system. The goal of classification algorithms is to predict the value of the output also for new cases, whose label value is not known.

Nevertheless, in many contexts the adoption of the so-called clear box methods is as well worthy of attention. Those methods are able to produce the predictive model in the form of intelligible rules in the IF-THEN form. This class of techniques should be conveniently used in the antifraud context, generally allowing a good level of integration with legacy systems.

Involved datasets:

- Accident data (place, damage, ...);
- Policy data (starting date, type, ...);
- Payment data (e.g., multiple payments to the same account);
- Existing rules;
- Data from publicly available datasets.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

- 30% increase in fraud detection

- 7% decrease of false positive

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

ML is already a mature technology; nevertheless, applications in real world contexts, in particular in more traditional sectors, is still an ongoing process. Some of the main obstacles preventing diffuse adoption of machine learning approaches:

- **Integration with legacy systems;**
- **Difficulties in the deployment of the solutions in real world scenarios and in integration with IT systems;**
- **Gap in skilled people able to effectively use this type of approaches.**
- **These are aspects that should be considered when planning any type of solution in this field by choosing technologies able to:**
- **Integrate and interface with already working IT systems;**
- **Reduce the need of skilled people (e.g., have visual interfaces and do not require programming).**

II.4 REGTECH

II.4.1 Automated Regulatory Risk Intelligence

Responsible:	CeADAR
Sector:	RegTech
Typology:	Nationally funded
State of the Art:	Under construction
Keywords:	Regulatory Risk Intelligence, Regulatory Monitoring, Horizon Scanning

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Today, large firms deploy significant financial resources and human capital to control regulatory risks. Regulatory rulebooks are updated continuously with firms using entirely manual processes to cope with a huge workload. As a result, personnel costs in compliance functions have greatly increased, with large firms now employing over 10,000 staff in control roles. Small firms face similar challenges but on a smaller scale.

The regulatory risk landscape changes in an unpredictable manner in response to global issues and individual enforcement actions, making it difficult for firms to monitor and assess the most relevant themes, business service lines, jurisdictions and regulatory areas in a timely manner. Recent progress in data science and artificial intelligence as applied to unstructured data streams

now brings the prospect of automated horizon scanning and evaluation of regulatory risks within reach.

The opportunity in this case is to provide new services based on targeted thematic assessment powered by natural language processing of regulatory documents, in order to direct manual resources and streamline filtering and assessment of the large volume of regulatory updates and news that appear daily. The potential impact of regulatory enforcement on financial companies is large and growing, but difficult to quantify in comparison to other forms of legal and corporate risk, such as credit or market risk. The ability to automatically categorise and quantify the kinds of potential impact described in these documents will be of great benefit to businesses in complying with increasingly complex and prominent requirements.

2. Methodology

Description of the use case and its methodology

Regulatory updates and the flow of regulatory attention as revealed in news and online discussion are available as open documents on the websites of regulators, media, and social media. The focus in this use case is to extract signals from these streams using natural language processing to identify features of interest, and machine learning to classify categories of interest and monitor for early identification of topical trends. Technology at this stage of commercialisation can enable services such as:

Classification of regulatory updates: Enforcement actions, announcements, or regulatory news can be flagged according to control area or business service line using automated document classification;

Topical trend analysis: Topic modelling the stream of regulatory updates can identify rising trends as they begin to increase in prominence over time in a collection of regulatory updates;

Modelling risk across categories: Statistical analysis of previous records of enforcement actions allows for quantification of risk by regulatory category, topic, and business service line.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

To facilitate text analysis, regulations, enforcements, announcements, and news relevant to regulatory risk will be collected from open online sources. Many of these sources facilitate automated access to the relevant documents. The focus of this use case is unstructured data, primarily text, along with relevant metadata such as jurisdiction, dates, regulatory authority, and the nature of the document, collected alongside the text.

Topic modelling (Latent Dirichlet Allocation) will be applied to identify clusters of terms corresponding to topic of interest. This is an unsupervised process, allowing for important new themes to arise from the incoming data stream in a bottom-up manner.

Document classification with appropriate supervised learning algorithms such as neural networks or random forest classifiers will allow for automated classification of documents by regulatory category or service line, as well as the identification of words and phrases that are associated with particular categories.

In both of these technologies, the natural language processing component includes named entity recognition and extraction of relevant multi-word expressions.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The solution will present a dashboard of incoming regulatory documents classified by their predicted regulatory and business service categories. These may be of use directly, or may be guided to domain experts for further analysis where the importance level of the category is high. Trends in the totals of these categories as well as emerging topics may be presented as trendlines on such a dashboard.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations and lessons learnt

The main foreseen challenges would include:

- The document and news sources need to be regularly maintained and kept up to date both with regulators and more general sources of news signal;
- The global nature of the regulatory landscape means linguistic expertise across a selection of important jurisdictions may be necessary to configure and evaluate the automated process;
- Tail risks mean that existing risk ontologies may not contain important new shifts in regulatory attention -- the unsupervised topic discovery process requires regular expert monitoring to evaluate new risks on the horizon.

II.4.2 Harmonised Compliance Protocol

Responsible:	Muinmos
Sector:	RegTech, financial services
Typology:	EU project
State of the Art:	Proof of concept / Under construction / Parts already tested / Parts on the market
Keywords:	Regulatory client onboarding, Regulatory monitoring, Regulatory Compliance

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Globally, financial institutions often automate their KYC/AML checks whilst relying on more traditional legal reading of compliance in client categorisation as well as suitability and appropriateness of financial products and services to determine client onboarding outcomes also on a cross-border basis. Many are not attuned to the fact that robust compliance goes beyond just KYC/AML checks. In an increasing cross-border context even within the EU framework, discrepancies and differences in regulatory frameworks make for highly costly legal endeavours and the timelines taken for client onboarding can stretch for weeks and months.

Keeping up-to-date and remaining compliant in a fluid regulatory landscape across multiple jurisdictions is effectively a nightmare for any financial institution no matter its size. Through Global Regulatory Implementation Protocol (GRIP), regulatory burdens can be resolved and regulatory oversight strengthened in national and cross-border settings. A harmonised protocol also protects all parties (regulators, financial institutions and clients) from mis-interpretation/mis-selling and prevent the opportunity of regulatory arbitrage. The proposed data sharing protocol outlined will enable further innovation within RegTech and assist regulators in doing their job in a

cross-border environment whilst maintaining registers which all regulators and financial institutions across the globe can rely upon.

2. Methodology

Description of the use case and its methodology

The risk is that the financial institution offers a service that it is not allowed to offer and thereby risk breaching the law unknowingly where it would be subjected to reversal of trades and fines which is very large and growing, furthermore putting the investor at risk of their financial wellbeing. Not a single financial institution has managed to secure against this without incurring enormous legal fees and huge number of cross-department man hours in handling this.

Through the use of GRIP financial institutions can with ease classify their client in accordance with applicable regulation, and determine whether a financial institution can on-board a client to trade with it in a particular service, product and country whilst performing appropriateness and suitability assessments. This will enable financial institutions to comply with global regulation and at the same time provide increased security and protection for the investors. Today, gathering this data is a huge manual task and often exacerbated with errors. The establishment of a protocol is to facilitate automatic exchange of this kind of information. To the extent that this protocol is devised as an independent technology, it can be used by anyone with a need to retrieve regulatory permission sets. The objective is to help facilitate a formal mechanism for data sharing among regulators and third-party providers. It could, for example, take the form of a communication protocol, an API, a standard, or a file type specification.

- Currently, regulatory permissions/ licenses granted to financial institutions are made available on regulators' website, some in excel sheets, word documents and similar, and in some countries they are simply not available;
- Currently, legislative provisions that determine the regulatory permissions/ licenses granted to financial institutions are made available on authorities' website, some in pdf documents and similar.

The purpose of creating GRIP is to create a common terminology to eliminate confusion in order to aid the extraction, processing, analysis and correction of all this information from all sources.

The solution will be leveraging Natural Language Processing and Deep Learning in order to classify documents and extract specific information automatically. The relevant documents/ mediums handling regulatory permissions and legislative provisions are extracted from all relevant sources and sent through GRIP for processing and analysis). The deep learning algorithm will be able to extract all the relevant information, order the data, have GRIP make the analysis on them, trigger any discrepancies etc., then send it out for use by the authorities and financial institutions through various platforms.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

Through the use of GRIP, regulators can instantly see discrepancies between their regulatory permissions granted to the financial institutions based on the legislative provisions and correct them before providing them to the financial institution. The same holds for legislative provisions where any inaccuracies triggered between the laws and regulations and their actual implementation is highlighted which can be corrected at the source before implementation.

Using Natural Language Processing, Machine Learning, Neural Nets and Python language, it will not only create accurate regulatory permissions for each financial institution in line with the legislative provisions prevailing at all times, but it will also assess their probability of having an impact on the practical outcomes for financial institutions, which in turn has an impact on the client onboarding and market. This can develop an AI capable of suggesting mappings from standard syntax and word combinations which relate to legal metadata and tags, and can also comprehend significant clause updates and their probability of legal ramifications.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

Formatting of GRIP - Terms of the key protocol and its implementation:

- A. Highlighting all relevant legislative provisions and their expected behaviour in the areas covered;
- B. Highlighting all the relevant regulatory permissions and their expected behaviour;
- C. Using various available Natural Language Processing techniques for extraction of the legislative provisions and regulatory permissions in various languages;
- D. Using machine learning, neural nets and python to extract, process and analyse existing laws and regulation in alignment with regulatory guidance using GRIP to fulfil the requirements of the relevant legislative provisions, triggering automatic inaccuracies to the relevant authority;
- E. Using Machine learning, neural nets and python to gather, process and analyse of the permissions/ licenses granted to financial institutions to fulfil the requirements of the relevant legislative provisions, triggering automatic inaccuracies to the relevant financial institutions who in turn will seek correction from the relevant regulatory authority;
- F. Any discrepancies triggered in D. and/or E once notified to the relevant parties as above, the relevant party will have the ability to correct at the source which will in turn create a harmonised protocol protecting all parties (regulators, financial institutions and clients) from mis-interpretation/mis-selling and prevent the opportunity of regulatory arbitrage.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

Lack of the data being available in digital form across the globe in the marketplace makes the retrieval of information very hard in terms of legislative provisions and regulatory permissions, which impacts processes and makes it difficult to create a global AI engine to support the entire financial market.

However, it is likely that the financial sector is moving more and more towards automation which for the authorities and regulators will require inputting their existing data sets into a digitized form. This will become the norm and will make the project of GRIP much more realistic on a global scale in terms of creating a harmonized protocol for use by the financial institutions and regulators for the benefit of enabling financial institutions to comply and for the protection of the consumer. Working with many regulators around the world in aiding the thinking around GRIP and many authorities and regulators are beginning to see that digitization in this space is a must to enable the next generation and move away from manual processes of updating laws, regulations and permissions to an actual system framed around a common protocol for every player in the market place to use.

The market is much more ready to embrace such solutions today compared to some years ago. Such an implementation would create a high-water mark for the regulatory landscape to meet the needs of the 21st century financial world.

II.5 OTHERS

II.5.1 Audio Experience from Investors Reports

Responsible:

ATC

Sector:	Others
Typology:	Industrial
State of the Art:	On the market
Keywords:	Automation, financial and business reports

1. Business Motivation

Brief description of the challenge and the addressed business pain, highlighting the innovation potential

Success in the Financial, Banking and Insurances sectors is associated closely with informed decision-making, typically, based upon analysis and business reports that extract meaning out of large amounts of data, such as financial reports, data from other news sources, government sources and the internet. In particular, there exist many firms that compile and successfully trade investor information in document form. Typically, such information is distributed in electronic format, especially in .pdf format, (portal document format). The size of these documents range from large to very large and they may incorporate complex media types, including text, pictures, diagrams and formulas.

It has been realised that the large size and difficulty in reading of such complex documents constitutes a disadvantage for trading them, as it is time consuming and can be difficult to make sense and navigate through them. Therefore, a way to overcome these issues is the 'AI Journalism', to transform long and complex documents into human listenable audio experiences. In this way, information is made more accessible as readers are turned into listeners, also, whilst on the move via mobile phones and applications, while enhancing the rate of information absorption.

2. Methodology

Description of the use case and its methodology

Typically, companies publish and trade investor information in the range of thousands of documents per year which are managed by a content management system. This solution is based on an AI/ML-based text-to-speech solution, which transformed documents in formats which lack sufficient metadata and rich mark-up languages into audio experiences to be accessed through a mobile application and a complimentary web app. The solution should not require human intervention during conversion, especially in removing unreadable content and coping with domain terminology, syntax, abbreviations, while the end result should sound natural, as if a human was speaking.

3. Technological Components

High-level description of the utilized technologies and the datasets involved

Published and traded documents in .pdf file formats lack rich metadata that may describe their contents, hence, text-to-speech conversion becomes extremely challenging, if not impossible, with conventional convertors. Indicatively, the mark-up language, as it stands, can only capture the presentation of the content and not descriptions of the conveyed information. Moreover, OCR'd produced .pdf documents that are images of scanned pages are not machine readable, or lack a consistent flow that ADOBE conversion tools can process. Likewise, some 3rd party tools for .pdf document generation do not generate well-formed mark-up elements, adding an additional difficulty in converting text-to-speech. In conclusion and due to the above challenges, among others, the only information, typically, captured at the metadata level in .pdf documents are the position of the item and the type of font, which are insufficient for the automatic conversion of .pdf documents into audio experiences.

4. (Expected) Outcomes

The (expected) outcomes of the solution, KPIs reached, effectiveness

The developed AI/ML text-to-speech solution 'understands' the content within a .pdf document regardless of poor metadata and poor structure that it may contain. This solution, through a multi-stage AI/ML process, is capable of:

- Identifying text and blocks of text within a .pdf document;
- Understand and remove non-readable content such as images, footers, headers, references, labels, tables, figures etc.;
- Understand the flow of text within a page;
- Joining split sentences together, e.g., when a sentence continues from one page to the next, or continues from one column to the next, through the use of **NLP (Natural Language Processing)**;
- Understand and replace abbreviations and shorthand with the full text;
- Dealing with domain specific syntax.

The solution has an F₁ score (accuracy) of 80% and higher, which provides a clean and consistent text which is then used to create an audio mp3 of the document.

5. Challenges and Recommendations

Challenges and problems foreseen/faced (technology maturity, business, implementation), recommendations, and lessons learnt

The main challenges relate to the availability of:

- Large enough corpus of documents for AI/ML algorithm training with all variations possible;
- Extensive set of domain specific abbreviations;
- Well and complete set of domain specific syntax.

Bibliography

Baitinger, E. and Papenbrock, J. (2017). Interconnectedness risk and active portfolio management, *Journal of Investment Strategies*, Risk.net, Volume 6, Number 2

Battiston, S., Delli Gatti, D., Gallegati, M, Greenwald, B., Stiglitz, J. E. (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control*, 36(8), pp. 1121-1141.

Billio, M, Getmansky, M, Lo, A: and Pellizzon, L. (2012) Econometric measures of systemic risk in the finance and insurance sector. *Journal of Financial Economics*.

Cruz, M.G. (2002). *Modelling, Measuring and Hedging Operational Risk*. Wiley Finance;

Diebold, F. and Yilmaz, K. (2014). On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms. *Journal of Econometrics*, 182(1), pp. 119-134;

European Commission (2018). *Fintech Action Plan: For a More Competitive and Innovative European Financial Sector*.

Financial Stability Board. (2017a). *Artificial intelligence and machine learning in financial services: Market developments and financial stability implications*.

Financial Stability Board. (2017b). *Financial Stability Implications for Fintech: Supervisory and Regulatory Issues that Merit Authorities' Attention*.

Foley, S., Karlsen, J.R. and Putnins, T.J. (2018). *Sex, drugs and bitcoin: How much illegal activity is financed through cryptocurrencies*. Working Paper;

Giudici, P. (2003). *Applied data mining*. Wiley, London.

Giudici, P. and Hadji-Misheva, B. (2019). *Network scoring models for P2P lending*. *Quality engineering*.

Giudici, P. and Polinesi, G. (2019). *Risk classification methods in Robot Advisory Platform*. *Journal of Network theory in finance*.

Giudici, P. and Abu-Hashish (2019). What determines bitcoin exchange prices: a network VAR approach. Finance Research letters.

Giudici, P., Spelta, A. (2016). Graphical network models for international financial flows. Journal of Business and Economic Statistics 34 (1), pp. 126-138;

Lauritzen (1996), Graphical models, Wiley,

Makrichoriti, P. and Moratis, G. (2016). BitCoin's Roller Coaster: Systemic Risk and Market Sentiment. Working Paper;

Mantegna (1999). The European Physical Journal B: Condensed Matter and Complex Systems, 1999, vol. 11, issue 1, 193-197

Papenbrock, J; Schwendner, P. (2015). Handling Risk On/Risk Off Dynamics with Correlation Regimes and Correlation Networks. Financial Markets and Portfolio Management, 29(2), pp.125-147

Scherer, M.U. (2016). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. Harvard Journal of Law and Technology, 29(2), pp.1-48

Tim Miller - Explanation in Artificial Intelligence: Insights from the Social Sciences (2018, Journal of Artificial Intelligence 267, 1-38)

Arrieta, A., Diaz-Rodriguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., e Herrera, F - Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI (2020, Journal of Information Fusion, Vol. 58, 82-115).

Ribeiro, M., Singh, S., e Guestrin, C. - Why Should I Trust You? Explaining the Predictions of Any Classifier (2018, International Conference on Knowledge Discovery and Data Mining).

Carletti, M., Masiero, C., Beghi, A., e Susto, G. Explainable Machine Learning in Industry 4.0: Evaluating Feature Importance in Anomaly Detection to Enable Root Cause Analysis (2019).

Davenport, T. - The AI Advantage: How to Put the Artificial Intelligence Revolution to Work (2018)

Bussman, Giudici, Marinelli, Papenbrock (2020). Explainable machine learning in credit risk management. Computational economics, 2020

Giudici, Raffinetti. Shapley Lorenz decompositions (2020) Expert systems with applications.

Agosto, Giudici, A poisson autoregressive model to understand COVID-19 contagion Risks, 8(3), 1-8, 77 (2020)

Giudici, P., Hadji-Misheva, B., Spelta, A. Network based credit risk models. Quality Engineering 32 (2), 199-211 (2020)

Avdjiev, S., Giudici, P. and Spelta, A. Measuring contagion risk in international banking. Journal of Financial Stability, 42, 36-51 (2019)

Ahelegbey, D., Giudici, P. and Hadji-Misheva, B. Latent factor models for credit scoring in P2P systems. Physica A: statistical mechanics and its applications, 522, 112-121 (2019)

Paola Cerchiello, Paolo Giudici, Giancarlo Nicola: Twitter data models for bank risk contagion. Neurocomputing, 264, 50-56 (2017)

Giudici, Abu-Hashish. What determines bitcoin exchange prices? A network VAR approach. Finance Research Letters vol. 28, 309-318 (2019)

Facchinetti, S., Giudici, P. and Osmetti, S. Cyber risk measurement with ordinal data. Statistical methods and applications 29 (1), 173-185 (2020)

Editorial Team and Contributors

Editors:

Marina Cugurra (GFT)

Vittorio Monferrino (GFT)

Marco Rotoloni (ABI LAB)

Barbara Cacciamani (ABI LAB)

Tomislav Duricic (Know-Center)

Contributors:

Nr.	Title	Contributors
1	Mission beyond Digital Finance	Marina Cugurra (GFT), Maurizio Ferraris (GFT)
2	Digital Finance Ecosystem	Tomislav Duricic (Know-Center)
2.1	Market and trend analysis	Tomislav Duricic (Know-Center)
2.2	The Challenge of Digital Transformation	Maurizio Ferraris (GFT), Tomislav Duricic (Know-Center)
2.3	Digital Transformation Areas in the Financial Sector	Maurizio Ferraris (GFT), Tomislav Duricic (Know-Center)
2.3.1	Digital Finance State of the Art Solutions	Maurizio Ferraris (GFT), Tomislav Duricic (Know-Center)
2.4	The impact of COVID 19	Miguel Saldaña (Inmark)
3	Evolving Digital Finance towards 2025	Vittorio Monferrino (GFT)
3.1	Responsible Data-Driven AI	Marco Rotoloni, Barbara Cacciamani (ABILAB)
3.1.1	Data for AI	Marco Rotoloni, Barbara Cacciamani (ABILAB)
3.1.2	Security and Privacy	Ilesh Dattani (ASSENTIAN)
3.1.3	The evolving Regulatory Framework fostering EU competitiveness	Marina Cugurra (GFT)
3.1.4	Ethical AI in the Financial Sector	Marina Cugurra (GFT)
3.2	Deployment of applications and real-life scenario recommendations	Alessandro Mamelli, Domenico Costantino (HPE), Jacopo Mauro (University of Southern Denmark).
3.3	Europe: the major AI-Fintech Hub of the world	Francisco, Estevan (INSOMNIA DIGITAL INNOVATION HUB)

4.1	Artificial Intelligence and Robotics in Digital Finance	Marco Rotoloni, Barbara Cacciamani (ABILAB)
4.1.1	Ongoing challenges	Marco Rotoloni, Barbara Cacciamani (ABILAB)
4.1.2	EXplainable Artificial Intelligence (XAI)	Paolo Giudici (Pavia University), Enrico Ferrari (Rulex)
4.2	Big Data and Analytics in Digital Finance	John Soldatos (Innov-Acts) Davide Dalle Carbonare (Engineering), Susanna Bonura (Engineering).
4.2.1	Trends and Challenges	John Soldatos (Innov-Acts)
4.2.2	Emerging Big Data Solutions and examples of interesting applications	John Soldatos (Innov-Acts)
5	Recommendations towards future Research & Innovation Agendas	Vittorio Monferrino (GFT)
I.1	Digital Native Banks	Alessandro Zerbetto (LINKS Foundation)
I.2	Decentralized finance and comparative analysis of emerging crosschain services	Roberto Moncada (LINKS Foundation), Enrico Ferro (LINKS Foundation)
I.3	Decentralized approaches to manage datasets used by AI and Big Data Systems	Nadia Fabrizio (Cefriel)
I.4	Network finance. A systemic and network approach designed to ease liquidation of trade receivables, optimise short-term credit allocation, and enhance and improve the efficiency of risk analysis models	Carlo Mancosu, Nicola Pirina, Roberto Spano, Cesare Ravaglia, Rinaldo Bonazzo, Andrea Zanda, Simone Demelas, Nico Lauridsen, Lorenzo Pinna, Massimo Cossu
II	Collection of use cases	Vittorio Monferrino (GFT)
II.1.1	Intelligent Document Processing – IDP	Julien Donnet (GFT)
II.1.2	Data Driven Decisions – 3D	Athina Ropodi (Wings ICT)
II.2.1	Blockchain in the Digital Asset Supply Chain	Oisin Boydell (CeADAR, University College Dublin), Edward McDonnell ((CeADAR, University College Dublin), Ricardo Simon Carbajo (CeADAR, University College Dublin)
II.2.2	DAOs in Decentralized Finance (DeFi)	Roberto Moncada (LINKS Foundation), Enrico Ferro (LINKS Foundation)

II.2.3	Network Finance	Carlo Mancosu, Nicola Pirina, Roberto Spano, Cesare Ravaglia, Rinaldo Bonazzo, Andrea Zanda, Simone Demelas, Nico Lauridsen, Lorenzo Pinna, Massimo Cossu
II.3.1	Personalized insurance products on IoT vehicles	Jose Gato (ATOS)
II.3.2	Fraud detection and prevention in car insurances	Enrico Ferrari (Rulex)
II.4.1	Automated Regulatory Risk Intelligence	David Lillis, CeADAR, University College Dublin, Ireland Edward McDonnell, CeADAR, University College Dublin, Ireland Ricardo Simon Carbajo, CeADAR, University College Dublin, Ireland
II.4.2	Harmonised Compliance Protocol	Remonda Kirketerp-Moller (Muinmos)
II.5.1	Audio Experience from Investors Reports	Antonis Ramfos (ATC)

Reviewers:

Davide delle Carbonare, Susanna Bonura (ENG)

Martin Serrano (NUIG)

Valerio Frascolla (Intel)

Mattia Trino (BDVA)

Ana Garcia Robles (BDVA)

In addition, special thanks to all BDVA/DAIRO members of the Task Force “Big Data and AI for Digital Finance”, who contributed to the content of this paper and actively collaborated in several dedicated Task Force Meetings and other BDVA events.

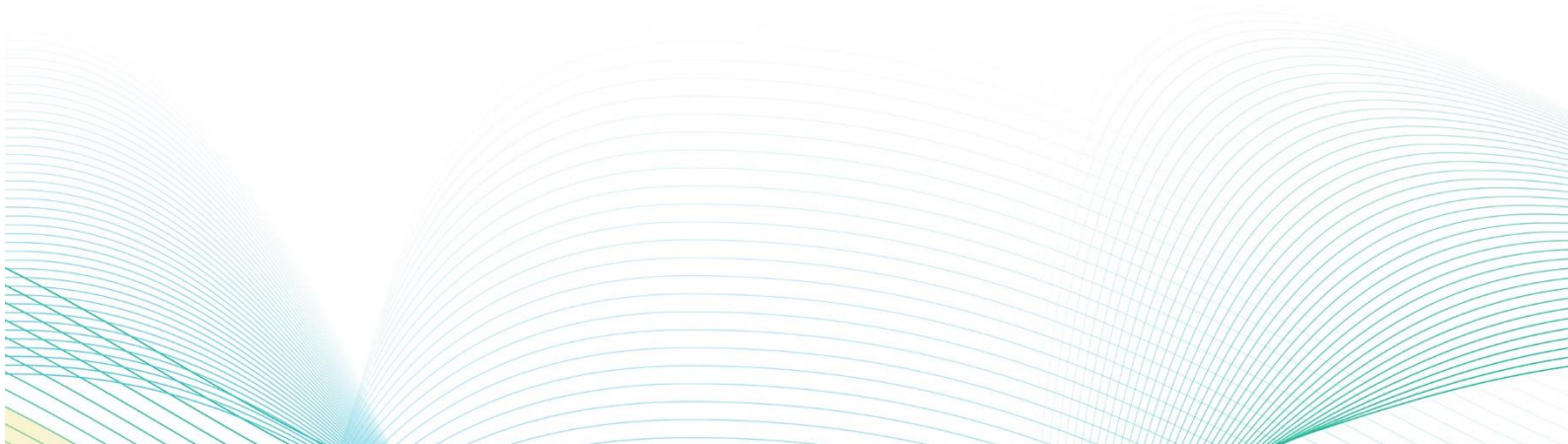
Note

This document should be referenced as follows: “AA.VV. (2021) AI and Big Data for the Financial Sector (v. 2022). Brussels. BDVA/DAIRO”

About BDVA

The Big Data Value Association – BDVA, (from 2021, DAIRO - Data, AI and Robotics aisbl), is an industry-driven international not-for-profit organisation with more than 230 members all over Europe and a well-balanced composition of large, small, and medium-sized industries as well as research and user organizations. BDVA focuses on enabling the digital transformation of the economy and society through Data and Artificial Intelligence by advancing in areas such as big data and AI technologies and services, data platforms and data spaces, Industrial AI, data-driven value creation, standardisation, and skills. BDVA has been the private side of the H2020 partnership Big Data Value PPP, it is a private member of the EuroHPC JU, it is also one of the founding members of the AI, Data and Robotics Partnership and a partner in the Data Spaces Business Alliance. BDVA is an open and inclusive community and is always eager to accept new members who share these ambitious objectives.

Contact for further information: info@core.bdva.eu





www.bdva.eu